

SUMMARIES

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 14, 2002, 18:11:26 : Search time 139.234 Seconds  
(without alignments)  
3202.482 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDQLPSPATIACHLDP.....RPKPKIOTRREPTPIHLS 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09782953/runat\_1122002.114428\_17330/app\_query.fasta.1.1173  
-DB=N Geneseq 101002 -OPMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953@cgn2\_1.1220@runat\_1122002.114428\_17330 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	597	24	Human MCIP associa
2	1002.5	95.7	2212	24	Human MCIP associa
3	957.5	91.4	2358	21	Lung cancer associ
4	920.5	87.8	597	24	Mouse MCIP associa
5	872.5	83.3	2331	24	Human MCIP associa
6	859.5	82.0	2348	24	Human DNA sequence
7	847.5	80.9	2173	24	Gene #3341 used to
8	844.5	80.6	2174	20	Human DSCR1 coding
9	749.5	71.5	599	24	Human MCIP associa
10	620.5	59.2	934	22	Nucleotide sequenc
11	620.5	59.2	3159	24	Human Down syndrom
12	613.5	58.5	534	24	Mouse MCIP associa
13	596.5	56.9	3184	24	Human MCIP associa
14	596.5	56.9	3184	24	Colon adenocarcino
15	596.5	56.9	3184	24	Lung cancer relate
16	596	56.9	828	24	Human MCIP associa
17	595	56.8	720	24	Human MCIP associa
18	539	51.4	615	24	Rat Down syndrome
19	509.5	48.6	1021	24	Human secreted pro
20	439	41.9	442	21	Drosophila melanog
21	379.5	36.2	1820	23	Drosophila melanog
22	379.5	36.2	12550	23	Human foetal liver
23	285.5	27.2	412	22	Human brain expres
24	285.5	27.2	412	22	Human foetal liver
25	285.5	27.2	412	22	Human brain expres
26	285.5	27.2	412	22	Probe #4382 for ge
27	285.5	27.2	412	22	Probe #4507 used t
28	285.5	27.2	412	22	Human genome-deriv
29	285.5	27.2	412	24	Human foetal cell
30	285.5	27.2	446	22	Human foetal liver
31	285.5	27.2	446	22	Probe #2064 for ge
32	285.5	27.2	446	22	Human brain expres
33	285.5	27.2	446	22	Human bone marrow
34	285.5	27.2	446	22	Probe #2081 for ge
35	285.5	27.2	446	22	Probe #2179 used t
36	285.5	27.2	446	22	Probe #2052 used t
37	285.5	27.2	446	22	Human genome-deriv
38	266.5	25.4	486	22	Human breast cell
39	266.5	25.4	486	22	Human foetal liver
40	266.5	25.4	486	22	Probe #809 for gen
41	266.5	25.4	486	22	Human brain expres
42	266.5	25.4	486	22	Human bone marrow
43	266.5	25.4	486	22	Probe #836 for gen
44	266.5	25.4	486	22	Probe #849 used to
45	266.5	25.4	486	22	Probe #816 used to

ALIGNMENTS

RESULT 1  
AAD30152  
ID AAD30152 standard; DNA; 597 BP.

XX	AAD30152;	
XX	17-MAY-2002 (first entry)	
XX	Human MCIP associated DNA #2.	
XX	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;	
XX	heart failure; cardiomyopathy; heart disease; human; gene; ds.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	1..597	
XX	/*tag= a	

FT /product= "Human MCIP associated protein #2"

XX W0200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

XX P-PSDB; AAE18911.

XX Screening for modulators of muscle calcineurin interacting protein

PT (MCIP) binding, expression or phosphorylation, useful for treating

PT cardiac hypertrophy or heart failure, comprises mixing MCIP,

PT calcineurin and a test compound -

XX Claim 95; Page 147-148; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or

XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin

XX may be used for treating cardiac hypertrophy and heart failure.

XX Antibodies to MCIP can be used in characterizing the MCIP content of

XX healthy and diseased tissues and subsequently for determining the

XX presence or absence of cardiomyopathy or as predictor of heart disease.

XX The present sequence is human MCIP associated DNA.

XX

SO Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;

Alignment Scores:

Pred. No.: 1.57e-109 Length: 597

Score: 1048.00 Matches: 198

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0

US-09-782-953-3 (1-198) x MAD30152 (1-597)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAThrIleAlaCysHisLeuAspPro 20

DB 1 ATCGAGAGAGCTGATGATTCGACAGACCTGCGACGCGCCATGCGCTGACCTGACCG 60

OY 21 ArgValAlaPheValAspGlyLeuGlyCysArgAlaIlePheGluSerLeuPheArgThrTyrAsp 40

DB 61 CCGGCTGCTGTGGACGGCTGTGGCGGGCCAAATTGAATCCCTCTTCAAGAACATATAC 120

OY 41 LysAspThrThrPheGlnTyrPheLysSerPheIleValArgIleAsnPheSerAsn 60

DB 121 AAGGACACACCTTCACGATTTTAAAGACTTCAACGTCGCGGATTAACCTTCAGCAAC 180

OY 61 ProLysSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyValGlu 80

DB 181 CCTTATCTGCGAGCCGATCCAGGCTGCGCTGCGACACCGAGCTTCTGGGGAAGGAA 240

OY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 241 ATGAAGTGTATTGTCACAGACTTACATAGGAAGTTCACACTGCTCCGCCCAT 300

OY 101 ProAspLysGlnPheLeuIleSerProAlaSerProProValGlyTyrLysGlnVal 120

DB 301 CCGGACAAACAGTCTCTATCTCCCTCCGGCTCTCTCTCCGCTGGAAACAGTA 360

OY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140

DB 361 GAAGATGCCACCCCGCTGTAAATTACATCTTTATATGCGCATCTCCAGCTGGCGCA 420

OY 141 GlyLysLeuTyrGluLeuHisAlaIleThrAspProThrProSerValValHisVal 160

DB 421 GGAGAGAACTATGAACTGATCAGCAGACAGACCCCTCCAGTGTGTGTCCACGTG 480

OY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluMetGluArgMetLysArgPro 180

DB 481 TGTGAGAGTACCAAGACATGAGAGAGAGAGAGATGAGACATGAAAGAGACC 540

OY 181 LysProValIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198

DB 541 AAGCCCAAAATCATCCAGACAGGAGCGGAGTACACACCGATCCACTTACG 594

RESULT 2

AAD30156

ID AAD30156 standard; DNA; 2212 BP.

XX

AC AAD30156;

XX

DT 17-MAY-2002 (first entry)

XX

DE Human MCIP associated DNA #2.

XX

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

XX heart failure; cardiomyopathy; heart disease; human; gene; ds.

OS Homo sapiens.

XX

FX Key Location/Qualifiers

FT CDS 25..618

FT /'tag= a

FT /product= "Human MCIP associated protein #2"

XX

PN W0200204491-A2.

XX

PD 17-JAN-2002.

XX

PF 06-JUL-2001; 2001WO-US21662.

XX

PR 07-JUL-2000; 2000US-216601P.

XX

PR 13-FEB-2001; 2001US-0782953.

XX

XX (TEXA ) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

DR P-PSDB; AAE18915.

XX

PT Screening for modulators of muscle calcineurin interacting protein

PT (MCIP) binding, expression or phosphorylation, useful for treating

PT cardiac hypertrophy or heart failure, comprises mixing MCIP,

PT calcineurin and a test compound -

XX

PS Example 1; Page 159-161; 174pp; English.

XX

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or

XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin

XX may be used for treating cardiac hypertrophy and heart failure.

XX Antibodies to MCIP can be used in characterizing the MCIP content of



CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC Note: this sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.

XX  
 SQ Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

Alignment Scores:  
 Pred. No.: 1,38e-103 Length: 2212  
 Score: 1002.50 Matches: 191  
 Percent Similarity: 97.94% Conservatives: 3  
 Best Local Similarity: 96.46% Mismatches: 3  
 Query Match: 95.66% Indels: 1  
 DB: 24 Gaps: 1

US-09-782-953-3 (1-198) x AAD30156 (1-2212)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
 DB 25 ATGAGGAGGTGGACCTTCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCG 84  
 QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 DB 85 CGCGTGTTCGTGGACGCGCTGTGCCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144  
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 DB 145 AAGGACATCACCTTTCAGTATTTAAGAGCTTCAACGAGTCAGATAAATTCACGCAAC 204  
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 DB 205 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 264  
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
 DB 265 ATGAAGTATATATTTCGTGACACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAT 324  
 QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProValGlyTyrLysGlnVal 120  
 DB 325 CCAGCAAGAGATTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTG 384  
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140  
 DB 385 GAAGATCGACCCAGTCATAAATATGATCTTATATGCTATCTCAAGCTGGGCGCA 444  
 QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
 DB 445 GGGGAAAAGTATGAATTCACGCGAGCTGCACACCACTCCAGCGTGGTGGTCCATGTA 504  
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180  
 DB 505 TGTGAGAGTGATCAAGAG---AAGGAGAGAGAGAGAAATGGAAAGANTGAGAGACCT 561  
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
 DB 562 AAGCAAAAATATCCAGACCGAGGCGGAGTACAGCGGATCCACCTCAGC 615

RESULT 3

AAF18328

ID AAF18328 standard; DNA; 2358 BP.

XX AC AAF18328;

XX 14-MAR-2001 (first entry)

DT Lung cancer associated polynucleotide sequence SEQ ID 347.

DE Human; lung cancer associated protein; neuroprotective; cytoskeletal;

KW cardioactive; immunomodulatory; muscular active; vulnary;

KW gastrointestinal; nephrotropic; antiinfective; gynecological;

KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

KW proliferative disorder; wound healing; infectious disease; ds.

XX Homo sapiens.

XX W0200055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX P-FSDB; AAB58452.

XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer  
 XX  
 PS Claim 1; Page 805-806; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective, cytoskeletal, cardioactive,  
 CC immunomodulatory, muscular active general; vulnary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.

XX SQ Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

Alignment Scores:

Pred. No.: 1,93e-98 Length: 2358

Score: 957.50 Matches: 188

Percent Similarity: 96.46% Conservatives: 3

Best Local Similarity: 94.95% Mismatches: 6

Query Match: 91.36% Indels: 2

DB: 21 Gaps: 1

US-09-782-953-3 (1-198) x AAF18328 (1-2358)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
 DB 114 ATGAGGAGGTGGACCTTCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCG 173  
 QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 DB 174 CGCGTGTTCGTGGACGCGCTGTGCCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 233  
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 DB 234 AAGGACATCACCTTTCAGTATTTAAGAGCTTCAACGAGTCAGATAAATTCACGCAAC 293  
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 DB 294 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 353

```

OY 81 MetlyseuTYrPhealInThleuHialleglyseSerHisleuAlaProProAsn 100
DB 354 ATGAAGTATATTTTGTCTCAACCTTACATAGAAAGCTCACCTGGCTCCGCA--AAT 412
OY 101 ProAspIysGlnPheLeuHialleSerProAlaSerProProValGlyTTPlySGlnVal 120
DB 413 CCGAGACAGAGAGTTTCTGATCTCCCTCCCGCTCCGSCAGTGGATGGAAACAGTG 472
OY 121 GluAspAlaThrProValIleAenTYrAspLeuLeuTYrAlaIleSerIlyseuGlyPro 140
DB 473 GAAGATCCGACCCGACATCAATATGATCTTATATGTCATCTCCAGCTGGGGCCA 532
OY 141 G1yGluIlyerTYrGluLeuHialaIaThAspProThrProSerValIValHisVal 160
DB 533 GGGGAAAGATGATGATGACAGGAGCAGTACACCTCCGCGGTGGTCCATGTA 592
OY 161 CysGluSerAspGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
DB 593 TGTAGAGTATCAAGG---AGGAGGAAGAAAGAGAAATGAAAGATGAGAGACCT 649
OY 181 LysProIlyleIleIleGlnThrArgArgProGluTYrThrProIleHisleuSer 198
DB 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACGCCGATCCACCTCAGC 703

RESULT 4
AAD30153
ID AAD30153 standard; DNA: 597 BP.
XX
AC AAD30153;
XX
DT 17-MAY-2002 (first entry)
XX
DE Mouse MCIP associated DNA #3.
XX
KW Muscle calcineurin interacting protein; MCIP, cardiac hypertrophy;
heart failure; cardiomyopathy; heart disease; mouse; gene; de.
XX
OS Mus musculus.
XX
PH Key Location/Qualifiers
FT 1..597
FT CDS /product= "Mouse MCIP associated protein #3"
XX
PN MO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
PA (ROTH/) ROTHERMEL B.
XX
PI Williams SR, Rothermel B;
XX
DR MPI; 2002-179698/23.
DR P-PsDB; AAE18912.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Discloure; Page 150-151; 174pp; English.
XX
CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate

```

```

CC transcription of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterizing the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is mouse MCIP associated DNA.
CC Note: This sequence has been described as human MCIP3 encoding DNA
CC in the specification, however the sequence seems to be a polynucleotide
CC encoding a MCIP associated protein.
XX
SQ Sequence 597 BP, 170 A; 156 C; 142 G; 129 T; 0 other;

Alignment Scores:
Pred. No.: 4,63e-95 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 24 Gaps: 1

US-09-782-953-3 (1-198) x AAD30153 (1-597)
OY 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisleuAspPro 20
DB 1 ATGAGTTTATAGGACTTTAGCTACAAATTTTAGCTCCCGATGTTGTGTGGCAACGAT 60
OY 21 ArgValPheValAspGlyLeuCYsArgAlaIlyPheGluSerLeuPheAlaGlyThrAsp 40
DB 61 GATGCTTCAAGCAAGAGTGAAGACCAAGGCCAAATTTAAATCCCTTCAGAACATATGAC 120
OY 41 LysAspThrThrPheGlnTYrPheIlyseSerPheIlyAspValArgIleAspPheSerAsn 60
DB 121 AAGAGACACACTTCGATATTTAAAGCTTCAAAGCTGCCGATAACTTCAGAAC 180
OY 61 ProIlySerAlaIleAspAlaArgLeuArgLeuHisIlyThrGluPheGluIlyGlu 80
DB 181 CCTTATCTGACGCCGATGACCGAGCTGCGGTGCACAAGACCGAGTTCCTGGGGAAGAA 240
OY 81 MetlyseuTYrPhealInThleuHialleglyseSerHisleuAlaProProAsn 100
DB 241 ATGAAGTGTATTTTCTCAACTTTACATAGAAAGTTCACCTGGCTCCGCCCAAT 300
OY 101 ProAspIysGlnPheLeuHialleSerProAlaSerProProValGlyTTPlySGlnVal 120
DB 301 CCGCAAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGAAACAGTA 360
OY 121 GluAspAlaThrProValIleAenTYrAspLeuLeuTYrAlaIleSerIlyseuGlyPro 140
DB 361 GAAGATGCCACCCCGTCATCAATTTAGATCTTTATATCCCATCTCCAGACTGGGGCCA 420
OY 141 G1yGluIlyerTYrGluLeuHialaIaThAspProThrProSerValIValHisVal 160
DB 421 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
OY 161 CysGluSerAspGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
DB 481 TGTAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAG 540
OY 181 LysProIlyleIleIleGlnThrArgArgProGluTYrThrProIleHisleuSer 198
DB 541 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACACCGATCCACCTTACC 594

RESULT 5
AAD30155
ID AAD30155 standard; DNA: 2331 BP.
XX
AC AAD30155;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated DNA #1.

```

XX Musclé calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX Homo sapiens.  
 OS  
 XX  
 Key Location/Qualifiers  
 FT 144..737  
 CDS /tag= a  
 FT /product= "Human MCIP associated protein #1"  
 FT  
 XX WO200204491-A2.  
 PN  
 XX  
 PD 17-JAN-2002.  
 XX  
 XX 06-JUL-2001; 2001WO-US21662.  
 PF  
 XX  
 XX 07-JUL-2000; 2000US-216601P.  
 PR  
 XX 13-FEB-2001; 2001US-0782953.  
 PR  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 XX Williams SR, Rothermel B;  
 PI  
 XX WPI; 2002-179698/23.  
 DR  
 XX P-PSDB; AAEL8914.  
 DR  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 XX Claim 72; Page 155-157; 174pp; English.  
 PS  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcriptions of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterising the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC Note: This sequence has been described as a promoter in claim 72 of  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.  
 XX  
 SQ Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8.41e-89 Length: 2331  
 Score: 872.50 Matches: 168  
 Percent Similarity: 93.09% Conservative: 7  
 Best Local Similarity: 89.36% Mismatches: 12  
 Query Match: 83.25% Indels: 1  
 DB: 24 Gaps: 1  
 US-09-782-953-3 (1-198) x AAD30155 (1-2331)  
 Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
 Db 174 AGCTCCCTGATTGCTGTGTCGCAACAGTAGTATCTTCAGCGAAAGTGAACCCAGGCC 233  
 Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
 Db 234 AATTTGAGTCCCTCTTTAGGACGATGACAGGACATCACCTTTTCAGTATTTTAAAGC 293  
 Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70

Db 294 TTCAAAACGAGTCAGATAAACTTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 353  
 Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 Db 354 CTGCATAAGACTGAGTTCTTCGGAAAGGAATGAAGTTATATTGTCACACCTTACAC 413  
 Qy 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerProPro 110  
 Db 414 ATAGGAAGCTCACACCTCGCTCCGCCAAATCCAGACAAGCAGTTCTTGATCTCCCTCCC 473  
 Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
 Db 474 GCCTCTCCGCCAGTGGGATGGAAACAAGTGGAAAGATGCGACCCAGTCATAAATATGAT 533  
 Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
 Db 534 CTCATTATATGCACTCCAGCTGGGCGCAGGGGNAAGTATGAATTGCACGCGGACT 593  
 Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
 Db 594 GACACCACTCCCGAGCTGGTGGTCCATGTATGTGAGAGTGATCAAGAG--AAGAGGAA 650  
 Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190  
 Db 651 GAAAGAGGAATGGAAAGAAATGAGGAGACCTAAGCCAAAATTTATCCAGACCAGGAGCGG 710  
 Qy 191 GluTyrThrProIleHisLeuSer 198  
 Db 711 GAGTACACCGCATCCACTCAGC 734  
 RESULT 6  
 AAS94805  
 ID AAS94805 standard; DNA; 2348 BP.  
 XX  
 AC AAS94805;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human DNA sequence #60 expressed during foam cell differentiation.  
 XX  
 KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200177389-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 04-APR-2001; 2001WO-US111128.  
 XX  
 PR 05-APR-2000; 2000US-195106P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
 PI Tai J;  
 XX  
 XX WPI; 2002-010925/01.  
 DR  
 XX Composition useful for diagnosis of conditions, disorders or diseases  
 PT associated with atherosclerosis, comprises several polynucleotides that  
 PT are differentially expressed in foam cell development -  
 XX  
 PS Claim 1; Page 112-113; 315pp; English.  
 XX  
 CC The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a  
 CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the

CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used  
 CC as PCR primers and probes. The polynucleotides of the invention are also  
 CC useful in gene therapy. AAS94746-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation.

XX Sequence 2348 BP, 638 A, 473 C, 550 G, 687 T, 0 other;

# Alignment Scores:

Pred. No.:	2,546-87	Length:	2348
Score:	859.50	Matches:	168
Percent Similarity:	92.59%	Conservative:	7
Best Local Similarity:	88.89%	Mismatches:	12
Query Match:	82.01%	Indels:	2
DB:	24	Gaps:	1

US-09-782-953-3 (1-198) x AAS94805 (1-2348)

```

QY 11 SerAlaThrIleAlaCysHisIleuAspProArgValPheValAspGlyLeuCyatGala 30
DB 181 AGCTCCCTGATTCCTGTGTGGCAACAGTATATCTTACAGAAAGTAAACAGAGGCC 240
QY 31 LysPheGluSerLeuPheArgThrTyraAspLysAspThrThrPheGlnTyrrPheLysSer 50
DB 241 AATTTGAGTCCCTTTTGGAGCGATGACAGACATCATCCTTCAAGTATTTTAAAGAC 300
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70
DB 301 TTCAAACGAGTACGAAATTAACCTTACAGCAACCCCTTCCGAGCAGATCCAGGCTCCAG 360
QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHis 90
DB 361 CTGCATTAACAGTACGATTTCTGGGAAAGAAATGAAGTTATATTGCTCAGACCTTACAC 420
QY 91 IleGlySerSerHisLeuAlaProAsnProAsnProAsnProAsnProAsnProAsnPro 110
DB 421 ATAGGAAGCTCAGCTGAGGCTCCGCAATCCAGCAACAGCTTCTATCTCCCTCC 480
QY 110 AlaSerProProValGlyTyrPheGlnValGluAspAlaThrProValIleAsnTyrAs 130
DB 481 CCCCTCTCCGCGAGTGGATGGAACAAAGTGAAGATGCGACCCCGATCAATGAATGA 540
QY 130 PheLeuTyrrAlaIleSerLysLeuGlyProGlyGlyTyrrGluLeuHisAlaAlaThr 150
DB 541 TCTCTTATATGCGCATCTCCAACTGGGGCCAGGGGAAAGTATGAATTCACGCGCAGC 600
QY 150 rAspProThrProSerValIleValHisValCysGluSerAspGlnGluAsnGluGlu 170
DB 601 TGACACCACTCCCAAGCGTGTGTGTCATATATGTAGAGTATCAAGG---AAGGAGA 657
QY 170 UGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 190
DB 658 AAAAAAGGAAATGGAAGAAATGAGAGACCTAAAGCAAAATTAATCCAGACGAGAGGCC 717
QY 190 OGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 198
DB 718 GGAGTACAGCGCGATCCACTCAGC 742

```

## RESULT 7

ABN96843  
 ID ABN96843 standard; DNA; 2173 BP.

AC ABN96843;

XX 13-AUG-2002 (first entry)

DE Gene #341 used to diagnose liver cancer.

KM Gene; liver cancer; dv; hepatocellular carcinoma; hepatocytotic;  
 KM metastatic liver tumour; cytostatic; expression profile; disease state;  
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.

```

XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Perez-Da-Silva S, Vockley JG;
XX DR WPI, 2002-426119/45.
XX PT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample.
XX PS Claim 1; SEQ ID NO 3341; 298bp; English.

```

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN95503-ABN97455 in a tissue sample. The method of the invention has hepatocytotic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pat\\_sequences](http://wipo.int/pub/published_pat_sequences).

XX SQ Sequence 2173 BP, 606 A, 431 C, 503 G, 633 T, 0 other;

## Alignment Scores:

Pred. No.:	5,246-86	Length:	2173
Score:	847.50	Matches:	162
Percent Similarity:	97.63%	Conservative:	3
Best Local Similarity:	95.86%	Mismatches:	3
Query Match:	80.87%	Indels:	1
DB:	24	Gaps:	1

US-09-782-953-3 (1-198) x ABN96843 (1-2173)

```

QY 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrThrPheGlnTyrrPheLys 49
DB 58 GCCAAATTTGAGTCCCTTTTGGAGCGATGACAGACATCACTTCTGATATTTTAAAG 117
QY 50 SerPheLysArgValIleGlnIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeu 69
DB 118 AGCTTCAACAGTACGAAATTAACCTTACAGCAACCCCTTCCGAGCAGATGCCAGCTC 177
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeu 89
DB 178 CAGCTGCAATAGACTGAGTCTTCTGGGAAAGAAATGAAGTATATTGCTCAGACCTTA 237
QY 90 HisIleGlySerSerHisLysLeuAlaProProAsnProAsnProAsnProAsnProAsnPro 109
DB 238 CACATAGGAAGCTCAGACCTGCGCTCCGCAATTCAGACAGACAGTTCGATCTCCCTT 297
QY 110 ProAlaSerProProValGlyTyrPheGlnValGluAspAlaThrProValIleAsnTyr 129
DB 298 CCCCTCTCTCCGCGAGGAGTGAACAAAGTGAAGATGCGACCCGATCAATGAATGA 357
QY 130 AspLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGlyTyrrGluLeuHisAlaAla 149

```

|||||  
Db 358 GATCTCTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTGCACGCG 417  
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluGlu 169  
Db 418 ACTGACACCACTCCACGCGTGGTCCATGTATGTGAGAGTATCAAGAG---AAGGAG 474  
Qy 170 GluGluGluMetGluArgMetLysArgProLysProLysIleGlnThrArgArg 189  
Db 475 GAGAAGAGGAATGGAAAGATGAGAGACCTTAGCCAAAATTTATCCAGACGAGG 534  
Qy 190 ProGluThrThrProLysHisLeuSer 198  
Db 535 CCGGAGTACAGCGGATCCACCTCAGC 561  
RESULT 8  
ID AAX01282  
AA01282 standard; cDNA to mRNA; 2174 BP.  
XX AAX01282;  
XX AC AAX01282;  
XX 09-APR-1999 (first entry)  
XX Human DSCR1 coding sequence:  
XX DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;  
XX Central Nervous System development; mental retardation; heart defect; ds.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 49..564  
XX /\*tag= a  
XX polyA\_signal 1541..1546  
XX /\*tag= b  
XX polyA\_signal 2132..2137  
XX /\*tag= c  
XX US5869318-A.  
XX 09-FEB-1999.  
XX 07-JUN-1996; 96US-0665040.  
XX 07-JUN-1995; 95ES-0001140.  
XX (PALL/) PALLEJA X E.  
XX Fuentes JJ, Palreja XE, Pritchard M;  
XX WPI: 1999-152781/13.  
XX P-FSDB; AAW73898.  
XX DNA encoding foetal brain proteins - believed to be associated with  
XX Down's syndrome  
XX Claim 1; Column 15-18; 19pp; English.  
XX This sequence is the Down's Syndrome critical region 1 (DSCR1) gene  
XX of the invention. The DSCR1 gene was found to be located in the  
XX q22.1-22.2 region of human chromosome 21. An increase in the transient  
XX expression of DSCR1 mRNA in the brains of young rats, compared to  
XX expression levels in the brains of adult rats, suggests an important role  
XX for DSCR1 during the development of the Central Nervous System (CNS), and  
XX that over expression of DSCR1 may be involved in pathogenic abnormalities  
XX of mental retardation and/or heart defects as found in Down's syndrome  
XX patients.  
XX SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;  
Alignment Scores:  
Pred. No.: 1.15e-85 Length: 2174  
Score: 844.50 Matches: 161

Percent Similarity: 97.63% Conservative: 4  
Best Local Similarity: 95.27% Mismatches: 3  
Query Match: 80.58% Indels: 1  
DB: 20 Gaps: 1  
US-09-782-953-3 (1-198) x AAX01282 (1-2174)  
Qy 30 AlaLysPheGluSerLeuPheArgThrTyzAspLysAspThrThrPheGlnTyzPheLys 49  
Db 58 GCCAAATTTGAGTCCCTCTTTAGGAGCGTATGACAAGGACATCACCTTTTCAGTATTTTAA 117  
Qy 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAAsAlaArgLeu 69  
Db 118 AGCTTTCAACAGAGTCAGAAATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177  
Qy 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyzPheAlaGlnThrLeu 89  
Db 178 CAGCTGCATTAAGACTGAGTCTTGGGAAAGGAAATGAAGTATATTTTGTCTCAGACCTTA 237  
Qy 90 HisIleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerPro 109  
Db 238 CACATAGGAAGCTCACACTGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCT 297  
Qy 110 ProLysSerProProValGlyTyrPlyGlnValGluAspAlaThrProValIleAsnTyr 129  
Db 298 CCCGCTCTCCGCCAGTGGGATGGAACAACATGGGAAGATGGACCCAGTCTATAACTAT 357  
Qy 130 AspLeuLeuTyzAlaIleSerLysLeuGlyProGlyGluLysTyzGluLeuHisAlaAla 149  
Db 358 GATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGATGAATTCACGACGCG 417  
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
Db 418 ACTGACACCACTCCACGCGTGGTCCATGTATGTGAGAGTATCAAGAG---AAGGAG 474  
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleGlnThrArgArg 189  
Db 475 GAGAAGAGGAATGGAAAGATGAGAGACCTTAGCCAAAATTTATCCAGACGAGG 534  
Qy 190 ProGluThrThrProLysHisLeuSer 198  
Db 535 CCGGAGTACAGCGGATCCACCTCAGC 561  
RESULT 9  
AAD30151  
ID AAD30151 standard; DNA; 599 BP.  
XX AAD30151;  
XX AC AAD30151;  
XX DT 17-MAY-2002 (first entry)  
XX DE Human MCIP associated DNA #1.  
XX KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
XX heart failure; cardiomyopathy; heart disease; human; ds.  
XX OS Homo sapiens.  
XX PN WO200204491-A2.  
XX PD 17-JAN-2002.  
XX PF 06-JUL-2001; 2001WO-US21662.  
XX PR 07-JUL-2000; 2000US-216601P.  
XX PR 13-FEB-2001; 2001US-0782953.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX PA (WILL/) WILLIAMS S R.  
XX PA (ROTH/) ROTHERMEL B.  
XX PI Williams SR, Rothermel B;  
XX

DR WPI: 2002-179698/23.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 XX Disclosure: Page 147; 174pp; English.  
 XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC  
 XX  
 XX Sequence 599 BP; 164 A; 150 C; 149 G; 136 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1,18e-75 Length: 599  
 Score: 749.50 Matches: 151  
 Percent Similarity: 90.12% Conservative: 4  
 Best Local Similarity: 87.79% Mismatches: 14  
 Query Match: 71.52% Indels: 4  
 DB: 24 Gaps: 1  
 US-09-782-953-3 (1-198) x AAD30151 (1-599)  
 QY 5 AspleuGlnAspleuPro-----SerAlaThrIleAlaCyHisLeuAspProArg 21  
 DB 83 GATTATAGGACCTTATGACATTTTACCTCCGATGCTTGCTGGCAACACATGAT 142  
 QY 22 ValPheValAspGlyLeuCyValPheGlyLeuSerLeuPheArgThrTyAspLys 41  
 DB 143 GTCCTTCAGGAAAGTAGACCCAGGGCCAAATTTGAATCCCTTCAGAACATATGACAAG 202  
 QY 42 AspThrThrPheGlnTyPheLysSerPheLysArgValArgIleAspPheSerAspPro 61  
 DB 203 GACACACACCTTCCAGTATTTTAAAGACTTCAACCGTCCGATTAACCTCGCAACCCC 262  
 QY 62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGlyLysGluMet 81  
 DB 263 TTATCTGACGCGGATGCCAGCTGGCGGCGCAACAGCCGACTTCTGGGAGAGAAATG 322  
 QY 82 LysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaPProPheAspPro 101  
 DB 333 AAGTGTATTTTCTCAGACTTTACACATAGGAGAGTTCACACTGCTCCGCCAAT-CCC 381  
 QY 102 AspLysGlnPheLeuIleSerProPheAlaSerProProValGlyTyrLysGlnValGlu 121  
 DB 382 GACAAACAGCTTCTATCTATCCCTCCGCCCTCTCTCCCGTTGGCTGGAAACAGTAGAA 441  
 QY 122 AspAlaThrProValIleAsnTyrrAspLeuLeuTyrrAlaIleSerLysLeuGlyProGly 141  
 DB 442 GATGCCACCCCGCTCATAAATTACGATCTTTATATGTCATCTCAAGCTGGGGCAGAGA 501  
 QY 142 GluLysTyrrGluLeuHisAlaAlaThrAspProThrProSerValValHisValCys 161  
 DB 502 GAGAAAGTATGAATCTGACAGCAGCAGCAGACCACTCCACAGTGTGTGTCACCTGTGT 561  
 QY 162 GluSerAspGlnGluAsnGluGluGluGluGlu 173  
 DB 562 GAGAGTACCAAGAGATGACGAGCAAGAGAGAG 597  
 RESULT 10  
 AAF25338  
 ID AAF25338 standard; cDNA; 934 BP.

XX  
 AC AAF25338;  
 XX  
 XX 30-APR-2001 (first entry)  
 DT  
 XX  
 XX Nucleotide sequence of a human detoxification protein.  
 DE  
 KW Human; detoxification protein; DEXT; cancer; leukemia; melanoma;  
 KW adenocarcinoma; autoimmune disorder; inflammatory disorder;  
 KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
 KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
 KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.  
 XX  
 OS Homo sapiens.  
 FH  
 FH Key Location/Qualifiers  
 FT CDS 20..724  
 FT /\*tag= a  
 FT /\*product= "detoxification protein"  
 FT sig\_peptide 20..100  
 FT /\*tag= b  
 XX  
 XX MO200104305-A2.  
 XX  
 XX 18-JAN-2001.  
 XX  
 XX 06-JUL-2000; 2000MC-US18509.  
 XX  
 XX 07-JUL-1999; 99US-0142678.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 P1 Tang YT, Yue H;  
 P1  
 XX  
 XX WPI: 2001-147193/15.  
 DR  
 XX P-PSDB: AAB31788.  
 XX  
 PT New human detoxification protein and polynucleotide, useful for  
 PT diagnosis, prevention and treatment of autoimmune/inflammatory  
 PT disorders and cell proliferative disorders including cancer -  
 XX  
 PS Claim 5; Page 79; 79pp; English.  
 XX  
 CC The present sequence encodes a human detoxification polypeptide (DEXT).  
 CC DEXT and its (ant)agonists are useful for preventing or treating  
 CC disorders associated with decreased or increased expression or activity  
 CC of DEXT. DEXT polypeptides are useful for screening compounds that  
 CC specifically bind to DEXT and for identifying (ant)agonists.  
 CC Diseases prevented, treated and diagnosed include cancers (e.g.  
 CC leukemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,  
 CC breast, kidney, liver, pancreas, prostate and uterus),  
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,  
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative  
 CC colitis), bacterial, fungal, parasitic infections and cell  
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,  
 CC cirrhosis and hepatitis). Anti-DEXT antibodies may be used as  
 CC antagonists, as a targeting or delivery mechanism for bringing  
 CC pharmaceutical agents into contact with cells or tissues expressing  
 CC DEXT and for diagnosis of DEXT-related disorders.  
 XX  
 SQ Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 9.44e-61 Length: 934  
 Score: 620.50 Matches: 121  
 Percent Similarity: 73.71% Conservative: 22  
 Best Local Similarity: 62.37% Mismatches: 42  
 Query Match: 59.21% Indels: 9  
 DB: 22 Gaps: 2  
 US-09-782-953-3 (1-198) x AAF25338 (1-934)  
 QY 5 AspleuGlnAspleuProSerAlaThrIleAlaCyHisLeuAspProArgValPheVal 24



[illegible]

Alignment Scores:	3,16e-60	Length:	594
Pred. No.:	613, 50	Matches:	119
Score:	75.00%	Conservative:	22
Percent Similarity:	63.30%	Mismatches:	38
Best Local Similarity:	58.54%	Indels:	9
Query Match:	24	Gaps:	2
DB:			
US-09-782-953-3 (1-198) x AAD30154 (1-594)			
QY	11 SerAlaThrIleAlaCyNH:EleuAaPProAArgValPheValAaRGIyLeuCyAArgAla	30	
DB	31 TCCACTCGTCGTCGCTGTGTGTGAAGTCTTTACCAATCGAGAGTTAAAGAA	90	
QY	31 LysPheGluSerLeuPheArgThrTyAaPlyAaPThrThrPheGlnTyPheLysSer	50	
DB	91 AAATTCCAGAGGACCTGTTCCGAGCCATGATGATGATGAGCTTCCAGCTGTTAAAGT	150	
QY	51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAaPAlaArgLeuArg	70	
DB	151 TTCACACGGGTTCGAATTAATTCAAGCATCCCAAACTCGAGCCGTCGCCGATGAG	210	
QY	71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeuHis	90	
DB	211 CTTCAATGACACTCACTTCAAGGAGGAGAGCACTAAACTTACTTCCGCCAGGTCAAGAC	270	
QY	91 IleGlySerSer-----HisLeuAlaProProAsnProAaPlyGlnPheLeu	106	
DB	271 CCAGAGACGATGGAGACAAACTGCATTTGGCACCTCCACAGCTCGCCAAACAGTTCTTC	330	
QY	107 IleSerProProAlaSerProProValGlyTyProLysGlnValGluAaPAlaThrProVal	126	
DB	331 ATCTACACCCCTTCACTCTCCCTCGTGGCGAAGCTTAACAGCATGCGACACCGATC	390	
QY	127 IleAsnTyAaPLeuLeuTyAlaAlaSerLysLeuGlyProGlyGlnTyTyGlnLeu	146	
DB	391 CTCACACTGACCTTCTTTAAGCTGTGTGGCCAACTAAGACCAAGAGAAATATGACCTG	456	
QY	147 HisAlaAlaThrAaPProThrProSerValValHisValCysGluSerAaPArgGln	166	
DB	451 CACGCTGGAACCTGAAGTCAACCCAGAGCGTCGTGTGATGTGTGACAGGACATGAG	516	
QY	167 AsnGluGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGln	186	
DB	511 GAGGAGGAGACCCAAAG-----ACTTCCCCCAAGCCAAATAATCATTCAG	555	
QY	187 ThrArgArgProGluTyTyThrPro	194	
DB	556 ACCGGCGCTCCGGGCTTGCCACC	579	
RESULT 13			
AAD30157	standard; DNA; 3184 BP.		
AC	AAD30157;		
DT	17-MAY-2002 (first entry)		
DE	Human MCIP associated DNA #3.		
KW	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;		
KW	heart failure; cardiomyopathy; heart disease; human; gene; ds.		
OS	Homo sapiens.		
XX			
XX	Key Location/Qualifiers		
XX	205..783		
XX	FT CDS		
XX	FT		
XX	FT		
XX	PN		
XX	WO200204491-A2.		









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1118.55 Seconds  
(without alignments)  
2866.843 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQPLSATIACHLDP.....RPKPKIIQTRPEYTPIHLS 198

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-Q/cgn2\_1/USPTO.spool/US09782953/runat 11122002 114429 17353/app\_query.fasta\_1.1173  
-DB=EST -QPMI=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 @CNC 1.1.1583 @runat 11122002 114429 17353 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1002.5	95.7	828	9	AL536447
2	1002.5	95.7	890	9	AL536447
3	989.5	94.4	867	14	BQ427531
4	974.5	93.0	937	12	BE795722
5	963.5	91.9	956	9	AL556803
6	956.5	91.3	923	9	AL551657
7	904	86.3	1158	14	BQ895506
8	903	86.2	520	14	BQ749142
9	885	84.4	1507	11	AK010696
10	872.5	83.3	837	9	AU124628
11	872.5	83.3	931	9	AL543576
12	872.5	83.3	939	9	AL546617
13	872.5	83.3	1041	13	BM450020
14	865.5	82.6	945	14	BQ278576
15	863.5	82.4	718	9	AU131040
16	859.5	82.0	753	13	BI148584
17	859.5	82.0	885	9	AL544755
18	847.5	80.9	1014	9	AL559594
19	845.5	80.7	906	9	AL544313
20	843.5	80.5	662	13	BI464521
21	842.5	80.4	875	12	EG574693
22	841.5	80.3	939	9	AL554686
23	834.5	79.6	1078	13	BM541636
24	831.5	79.3	740	13	BI463566
25	827.5	79.0	950	12	BF782633
26	820	78.2	710	10	BB617325
27	812.5	77.5	711	13	EG915568
28	812.5	77.5	769	13	BI767955
29	811.5	77.4	811	12	EG570239
30	809.5	77.2	665	13	BI219142
31	801	76.4	980	12	BQ296537
32	797.5	76.1	905	14	BQ732056
33	790	75.4	827	13	BI102432
34	789.5	75.3	710	12	EG475986
35	784.5	74.9	689	9	AL652662
36	782	74.6	727	13	BI327875
37	776.5	74.1	642	14	BQ520533
38	769.5	73.4	898	14	BQ733563
39	763.5	72.9	751	9	AL576189
40	759.5	72.5	664	13	BQ061844
41	757.5	72.3	526	10	AW957479
42	749.5	71.5	599	9	AA200984
43	747.5	71.3	627	13	BI859506
44	731.5	69.8	1047	12	BQ287042
45	720.5	68.8	599	13	BQ064020

ALIGNMENTS

RESULT 1  
AL536447

LOCUS  
DEFINITION

AL536447 LTI FL013 FBRn1 Homo sapiens

prime, mRNA

828 bp

linear

EST 13-FEB-2001

CS0DF015YF21 5

Accession

AL536447

Version

EST

Keywords

GI:12799940

Source

human.

Organism

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Reference

1 (bases 1 to 828)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: sege@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers  
source 1..828

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODP0157F21"  
/clone\_1ib="LTI FL013\_FBrn1"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"

/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 194 a 235 c 261 g 137 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1.08e-102 Length: 828  
Score: 1002.50 Matches: 191  
Percent Similarity: 97.98% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 3  
Query Match: 95.66% Indels: 1  
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL536447 (1-828)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCySHisLeuAspPro 20  
Db 216 ATGGAGGAGGTGACCTGGAGGACCTGCCACGCACATCGCTGACCTGGAGCCG 275  
OY 21 ArgVal1PheVal1AspGlyLeuCyArgAla1Ala1PheGluSerLeuPheArgThrTyrAsp 40  
Db 276 CCGCGTGTCTGAGAGCGGCTGTGCGGCGCAATTGAGTCCCTCTTGGACGATGAC 335  
OY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal1ArgIleAsnPheSerAsn 60  
Db 336 AAGGACATCACCTTCAGATTTTAAGCTTCAACGAGTCAGATTAACCTTCAGCAAC 395  
OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 396 CCGTCTCCGAGAGAGTCCAGGCTCCAGTCGATTAACGATTTCTGGAAAGGAA 455  
OY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100  
Db 456 ATGAAGTTATATTTTGCTCAGACCTTACATAGGAAGCTTCACCTGCTCCGCAAT 515  
OY 101 ProAspLysGlnPheLeuLysSerProAlaSerProProVal1GlyTyrPheGlnVal 120  
Db 516 CCAGACAGACAGCTTCTGATCTCCCTCCGCTCCGCGCAGTGGAGTGAACAAGTG 575  
OY 121 GluAspAlaThrProVal1IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 576 GAAGATGCGAGCCCGAGTCAATAATGATCTTATATGTCATCTCCAAAGCTGGGGCCA 635  
OY 141 GlyLysLysTyrGluLeuHisAla1Ala1ThrAspProThrProSerVal1Val1HisVal 160  
Db 636 GGGGAAAGTATGATTTGACCCAGCAGCTACACCACTCCAGCGGTGGTGCATGTA 695  
OY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
Db 696 TGTGAGATGATCAAGAG---AAGGAGGAAGAAAGAAATGAAGAAATGAGAGACTT 752

OY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
Db 753 AAGCCAAAATATTCAGACCGAGGCGGAGTACAGCCGATCACCCTCAGC 806

## RESULT 2

AL538796 890 bp mRNA linear EST 16-FEB-2001  
LOCUS AL538796 LTI FL013\_FBrn1 Homo sapiens cDNA clone CSODP017Y122 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL538796  
VERSION AL538796.1 GI:12867423  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 890)  
AUTHORS Li, W.B., Gruber, C., Jeejee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: sege@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers  
source 1..890

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODP017Y122"  
/clone\_1ib="LTI FL013\_FBrn1"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"

/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 207 a 246 c 280 g 157 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.17e-102 Length: 890  
Score: 1002.50 Matches: 191  
Percent Similarity: 97.98% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 3  
Query Match: 95.66% Indels: 1  
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL538796 (1-890)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCySHisLeuAspPro 20  
Db 205 ATGGAGGAGGTGACCTGGAGGACCTGCCACGCACATCGCTGACCTGGAGCCG 264  
OY 21 ArgVal1PheVal1AspGlyLeuCyArgAla1Ala1PheGluSerLeuPheArgThrTyrAsp 40  
Db 265 CCGCGTGTCTGAGAGCGGCTGTGCGGCGCAATTGAGTCCCTCTTAAAGAGTATGAC 324  
OY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal1ArgIleAsnPheSerAsn 60  
Db 325 AAGGACATCACCTTCTGATTTTAAGAGCTTCAACGAGTCAAGATTAACCTTCAGCAAC 384  
OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 385 CCGTCTCCGAGAGAGTCCAGGCTCAGTAAAGCTGATTTCTGGGAAAGGAA 444  
OY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100

```

Db 445 ATGAAGTTATATTTTCTCAGACCTTACACATAGAGCTCACCTGGCTCCGCCAAT 504
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlysGlnVal 120
Db 505 CCAGACAAGCAGTTTCTGATCTCTCCCTCCCGCTCTCCGCGAGTGGATGGAAACAAGTG 564
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 565 GAAGATGCGACCCAGCTCAATAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 624
Qy 141 GlyLysLysTyrGlnLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 625 GGGGAAAGTATGAATGACGAGCGAGCTGACACCACTCCGAGCGTGGTGCATGTA 684
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 685 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGAGAAATGGAAAGATGAGGAGACT 741
Qy 181 LysProLysIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 742 AAGCAAAATTTATCCAGACCGAGCGGAGTACAGCGCGATCCACTCAGC 795

RESULT 3
BQ427531 867 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7914139 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156873
5', mRNA sequence.
ACCESSION BQ427531
VERSION BQ427531.1 GI:21166607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13501 row: n column: 10
High quality sequence stop: 626.
FEATURES
source
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6156873"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 204 a 241 c 267 g 155 t
ORIGIN
Alignment Scores: 3.3e-101 Length: 867
Pred. No.: 989.50 Matches: 191
Score: 97.49% Conservative: 3
Percent Similarity: 95.98% Mismatches: 3
Best Local Similarity: 94.42% Indels: 2
Query Match: 14 Gaps: 1
DB:
US-09-782-953-3 (1-198) x BQ427531 (1-867)

```

```

Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 155 ATGAGAGAGGTGGACCTGCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCCG 214
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 215 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 274
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 275 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAGAGCTCAGATAAATTTTCAGCAAC 334
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 335 CCCTTCTCCGCGAGCAGATGCCAGCTCCAGCTGCATAAGACTGAGTTCTTGGGAAAGGAA 394
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 395 ATGAAGTTATATTTTGTCTCAGACCTTACATAGTAAGAGCTCACACCTGGCTCCGCCAAT 454
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlysGlnVal 120
Db 455 CGAGCAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTG 514
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 515 GAAGATGCGACCCAGCTCAATAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 574
Qy 141 GlyLysLysTyrGlnLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 575 GGGGAAAGTATGAATGACGCGAGCGAGTACACCACTCCGAGCGTGGTGCATGTA 634
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 635 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGAGAAATGGAAAGATGAGGAGACT 691
Qy 181 LysProLysIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 692 AAGCAAAATTTATCCAGACCGAGCGGAGTACAGCGCGATCCACTCAGC 746

RESULT 4
BQ795722 937 bp mRNA linear EST 20-SEP-2000
LOCUS 601590730F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3944959 5',
mRNA sequence.
ACCESSION BQ795722
VERSION BQ795722.1 GI:10216920
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 937)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM802 row: k column: 08
High quality sequence stop: 812.
FEATURES
source
1..937
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3944959"
/clone_lib="NIH_MGC_7"

```

```

/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 215 a 244 c 286 g 192 t

ORIGIN

Alignment Scores:

Pred. No.:	1,76e-99	Length:	937
Score:	974.50	Matches:	190
Percent Similarity:	97.47%	Conservative:	3
Best Local Similarity:	95.96%	Mismatches:	4
Query Match:	92.99%	Indels:	2
DB:	12	Gaps:	1

US-09-782-953-3 (1-198) x BE795722 (1-937)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAcysHisLeuAspPro 20

DB 37 ATGAGAGAGGTGACCTCGCAGGACCTGCCAGGCGCCACATCCCTGTCACCTGACCCG 96

OY 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyraP 40

DB 97 CGCGTGTTCGTGACCGCGCTGCGGCCCAATTGAGTCCCTCTTTAGAGATATGAC 156

OY 41 LysAspThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 157 AAGGACATCACTCTTACATGATTTTAAAGCTTAAACGATCGAATTAACCTTCAGCAAC 216

OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 217 CCTCTCTCCGACGACATGCCAGCTCCAGCTGCATTAACACTGATTTCTGGAAAGGAA 276

OY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 277 ATGAAGTTATATTGCTCTAGACCTTACACATAGAAAGTCAACCTGCTCCGCCAAT 336

OY 101 ProAspLysGlnPheLeuLysSerProProlAserProProValGlyTyrrLysGlnVal 120

DB 337 CCAGAAACACAGTTTGTATCTCCCTCCGCTCTCCGCCAGTGGATGAAACAAAGTG 396

OY 121 GluAspAlaThrProValIleAsnTyraPheLeuLeuTyraIleSerLysLeuGlyPro 140

DB 397 GAAGATGCCACCCCACTGCTAAACTATGATCTTATATGCACTCCCAAGCTGGGGCA 456

OY 141 GlyLysLysTyrrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160

DB 457 GGGAA-AGATTATGATATGACGACGACGACTGACACCACTCCAGGCTGGTGTTCATGTA 515

OY 161 CyArgLysSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 516 TGTGAGAGTATCAAGAG---AAGAGAGAAAGAAAGAAATGAAATGAAAGAGAACT 572

OY 181 LysPheLysGlyIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198

DB 573 AAGCCAAAATATATCCAGACCCAGAGGCGGAGTACACGCCGATCCACTCAGC 626

RESULT 5

AL556803

LOCUS

DEFINITION

AL556803 LTI\_NFL006.PL2 Homo sapiens cDNA clone CSDBK012YD10 5

prime, mRNA sequence.

ACCESSION

AL556803

VERSION

AL556803.1 GI:12899815

KEYWORDS

EST.

SOURCE

human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 956)

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT

Genoscope - Centre National de Sequencage

Genoscope - BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..956

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CSDBK012YD10"

/clone\_id="LTI\_NFL006.PL2"

/tissue\_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and vector. Library was normalized. Library was constructed by a division of invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 214 a 273 c 301 g 158 t 10 others

ORIGIN

Alignment Scores:

Pred. No.:	3.12e-98	Length:	956
Score:	963.50	Matches:	187
Percent Similarity:	95.96%	Conservative:	3
Best Local Similarity:	94.44%	Mismatches:	7
Query Match:	91.94%	Indels:	1
DB:	9	Gaps:	1

US-09-782-953-3 (1-198) x AL556803 (1-956)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAcysHisLeuAspPro 20

DB 289 ATGAGAGAGGTGACCTCGCAGGACCTGCCAGGCGCCACATCCCTGTCACCTGACCCG 348

OY 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyraP 40

DB 349 CGCGTGTTCGTGACCGCGCTGCGGCCCAATTGAGTCCCTCTTTAGAGATATGAC 408

OY 41 LysAspThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 409 AAGACATCACTCTTACATGATTTTAAAGCTTAAACGATCGAATTAACCTTCAGCAAC 468

OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 469 CCTCTCTCCGACGACATGCCAGCTCCAGCTGCATTAAGAGTGTTCGGAAAGGAA 528

OY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 529 ATGAAGTTATATTGCTGACACTTACACATAGAAAGTCAACCTGCTCCGCCAAT 588

OY 101 ProAspLysGlnPheLeuLysSerProProlAserProProValGlyTyrrLysGlnVal 120

DB 589 CCAGACAAGCAAGTTTGTATCTCCCTCCGCTCTCCGCCAGTGGAGGAAACAAAGTG 648

OY 121 GluAspAlaThrProValIleAsnTyraPheLeuLeuTyraIleSerLysLeuGlyPro 140

DB 649 GAAGATGCCACCCCACTGCTAAACTATGATCTTATATGCACTCCCAAGCTGGGGCA 708

OY 141 GlyLysLysTyrrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160

DB 709 GGGGAAAAGTATGATATGACGACGACGACTGACACCACTCCCAAGCTGGTGTTCATGTA 768





Best Local Similarity: 92.02% Mismatches: 10  
 Query Match: 86.26% Indels: 0  
 DB: 14 Gaps: 0

US-09-782-953-3 (1-198) x BQ895506 (1-1158)

QY 11 SerAlaThrIleAlaCyHisLeuAspProArgValPheValAspGlyLeuCyArgAla 30  
 DB 139 AGCTCCGATTTGCTTGTGGCAACATGATGCTTCAGCCAGCAAGTGAAGACCGGCC 198

QY 31 LysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyrrPheLysSer 50  
 DB 199 AATTGATGATCCCTCTTCAGAACATATGACAGACACACCTTCAGATTTTAAGAGC 258

QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70  
 DB 259 TTCACACCTGCTCCGATTAACCTTCAGCAACCCCTTATCTCAGCCGATGCGAGCTGCGG 318

QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHis 90  
 DB 319 CTGCACAAAGCCGAGTCTCGGGAGAGAAATGAAGTTGTAATTTGCTCAGACTTTACAC 378

QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110  
 DB 379 ATAGGAAGTTCAACCTGGCTCCGCCCAATCCCAACAAAGTTCTCATCTCCCTCCG 438

QY 111 AlaSerProProValGlyTyrrGlnValGluAspAlaThrProValIleAsnTyraAsp 130  
 DB 439 GCCTCTCTCTCCCTGGCGAAGTAGAAGATGCCACCCCTCATTAATTAAGAT 498

QY 131 LeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaThr 150  
 DB 499 CTTTATATGCCATCTCCAGAGCTGGGCCGAGAGAGATATCACTGATGACGGCACA 558

QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170  
 DB 559 GACACCACTCCAGTGTGGTGTCCAGTGTGAGAGTGAACCAAGATGAGAGAGAG 618

QY 171 GluGluGluMetGluArgMetLysArgProLysProLysAlaIleGlnThrArgArgPro 190  
 DB 619 CAGGAAGATGAGAGAGATGAAGAGACCCCAATCATTCAGACAGAGAGACCG 678

QY 191 GluTyrrThrProIleHisLeuSer 198  
 DB 679 GAGTACACACCCATCCACTCAGC 702

RESULT 8  
 BQ749142 520 bp mRNA linear EST 17-JUN-2002  
 LOCUS UI-M-PD0-BYd-C-04-0-UI.r1 NIH BMAP\_PDO Mus musculus cDNA clone  
 DEFINITION IMAGE:5716539 5', mRNA sequence.  
 ACCESSION BQ749142  
 VERSION BQ749142.1 GI:21895929  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Eucelostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Dr. James Linn, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLN at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

FEATURES Seq primer: PYX-5.  
 Location/Qualifiers  
 source 1..520  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="MIMAG:5716539"  
 /clone\_1ib="NIH\_BMAP\_PDO"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 12.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /note="Organ: brain; Vector: pYX-AecI, Site 1: EcoR I, Site 2: Not I. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-AecI vector. The library tag is TAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 137 a 147 c 131 g 105 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,03e-91 Length: 520  
 Score: 903.00 Matches: 171  
 Percent Similarity: 99.42% Conservative: 0  
 Best Local Similarity: 99.42% Mismatches: 0  
 Query Match: 86.16% Indels: 0  
 DB: 14 Gaps: 0

US-09-782-953-3 (1-198) x BQ749142 (1-520)

QY 8 AspLeuProSerAlaThrIleAlaCyHisLeuAspProArgValPheValAspGlyLeu 27  
 DB 3 GACCTGCCAGAGCCGACCATGCTGCTCCACCTGACCCGCGCTGTCGAGCGGCTG 62

QY 28 CysArgAlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyrr 47  
 DB 63 TGCCGGCCCAATTGATCCCTTCAGAACATATGACAGACACACCTTCAGAT 122

QY 48 PheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAla 67  
 DB 123 TTTAAGAGCTTCAACGTGTCCGATTAACCTTCAGCAACCCCTTATCTCAGCCGATG 182

QY 68 ArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGln 87  
 DB 183 AGGCTGGCGCTGCACAAAGCCAGATCTCGGGAGAGAAATGAAGTGTATTTGCTCAG 242

QY 88 ThrLeuHisIleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIle 107  
 DB 243 ACTTACACATAGAGATTCAACCTGCTCCGCCCAATCCGACAAACAGTTCCTCATC 302

QY 108 SerProProAlaSerProProValGlyTyrrGlnValGluAspAlaThrProValIle 127  
 DB 303 TCCCTCCGCGCTCTCTCCGTGGTGGAAACAAATGAAGATGCCACCCCGTCATA 362

QY 128 AsnTyraAspLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHis 147  
 DB 363 AATTAGATCTTTATATGCCATCTCCAGCTGGGCCGAGAGAGATGATGAATGAT 422

QY 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167  
 DB 423 GCAGCAGACAGACCACTCCGAGTGTGGTGTCCAGTGTGTGAAGTGAACCAAGAGAT 482

QY 168 GluGluGluGluGluMetGluArgMetLysArg 179





BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

# FEATURES

Source

1. .931  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODI029XJ21"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t

## ORIGIN

Alignment Scores:  
Pred. No.: 5.39e-88 Length: 931  
Score: 872.50 Matches: 168  
Percent Similarity: 93.09% Conservativity: 7  
Best Local Similarity: 89.36% Mismatches: 12  
Query Match: 83.25% Indels: 1  
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL543576 (1-931)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
Db 80 AGCTCCCTGATTCCTGCTGGCAACAGTGATATCTTCAGCAAGTGAACAGGGCC 139  
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
Db 140 AAATTTCAGTCCCTCTTTAGGAGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 199  
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
Db 200 TTCAACAGCAGTCAGATAAATTTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 259  
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
Db 260 CTGCATGAAGCTGAGTTCTTCGGGAAGAAATGAAGTATATTTTGTCTCAGACCTTACAC 319  
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110  
Db 320 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCACAGACAGATTTCTGATCTCCCTCC 379  
Qy 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
Db 380 GCCTCTCCGCGAGTGGATGGAACAAGTGGAAAGATGCGACCCAGTCATAAATATGAT 439  
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysLysTyrGluLeuHisAlaAlaThr 150  
Db 440 CTCATTATGTCATCTCAAGCTGGGCGAGGAAAGATGAATGATGTCAGCGAGCACT 499  
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
Db 500 GACACCACTCCAGCGTGGTCCATGATGTCAGAGTATCAAGAG---AAGGAGAA 556  
Qy 171 GluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190  
Db 557 GAAGAGGAAATGGAAGAATCAGGAGACCTAAGCCAAAAATATCCAGCAGGAGGCCA 616  
Qy 191 GluTyrThrProIleHisLeuSer 198  
Db 617 GAGTACAGCGCGATCCACCTCAGC 640

RESULT 12

AL546617  
LOCUS

DEFINITION

AL546617 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CSODI029XJ21 5

prime, mRNA sequence.

ACCESSION

AL546617

VERSION

AL546617.1 GI:12879906

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 939)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .939

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CSODI029XJ21"

/clone\_lib="LTI\_NFL006\_PL2"

/tissue\_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 263 a 219 c 238 g 218 t

ORIGIN

1 others

Alignment Scores:

Pred. No.: 5.44e-88 Length: 939

Score: 872.50 Matches: 168

Percent Similarity: 93.09% Conservativity: 7

Best Local Similarity: 89.36% Mismatches: 12

Query Match: 83.25% Indels: 1

DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL546617 (1-939)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30

Db 83 AGCTCCCTGATTCCTGCTGGCAACAGTGATATCTTCAGCAAGTGAACAGGGCC 142

Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50

Db 143 AAATTTCAGTCCCTCTTTAGGAGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 202

Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70

Db 203 TTCAACAGCAGTCAGATAAATTTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 262

Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90

Db 263 CTGCATGAAGCTGAGTTCTTCGGGAAGAAATGAAGTATATTTTGTCTCAGACCTTACAC 322

Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110

Db 323 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCACAGACAGTTCCTGATCTCCCTCC 382

Qy 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130

Db 383 GCCTCTCCGCGAGTGGATGGAACAAGTGGAAAGATGCGACCCAGTCATAAATATGAT 442

```

Oy 131 LeuLeuTyraIaIleSerLyLeuGlyProGlyGluLyTyGluLeuHlaIaIaIaThr 150
Db 443 CTCTTATATGCTCCATCTCCAAAGCTGGGGCCAGGGAAGATATGATATGACGACGACCT 502
Oy 151 AspProThrProSerValValIaIaIaValCyAGLuserAspGlnGluangluGluGlu 170
Db 503 GACACCACTCCCGCTGGTGGTCCATGTATGTGAAGATCATCAAGAG--AAGAGAGAA 559
Oy 171 GluGluGluMeGluAArgMeLyAsArgProLyProLyLeIleIleGlnThArArArgPro 190
Db 560 GAAGAGAAATGAAAGATGAGAGAGACCTTAAGCCAAAATTATTCAGACGAGAGAGCCG 619
Oy 191 GluTyThrProIleHlaIeuser 198
Db 620 GAGTACACGCCGATCCACCTCAGC 643

RESULT 13
BM450020 1041 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
DEFINITION 5', mRNA sequence.
ACCESSION BM450020
VERSION BM450020.1 GI:18499060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12204 row: d column: 10
High quality sequence stop: 643.
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528025"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies"

BASE COUNT 282 a 248 c 274 g 236 t 1 others
ORIGIN

Alignment Scores:
Pred. NO.: 6.1e-88 Length: 1041
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.35% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 13 Gaps: 1

US-09-782-953-3 (1-198) x BM450020 (1-1041)
Oy 11 SerAlaThrTAlaIaCyHlaIeuserProArArgValPheValAspGlyLeuCyArArgAla 30
Db 114 AGCTCCGATGCTCGTGGTGGCAACAGTATATCTTCAGCGAAGTGAACAGGAGCC 173
Oy 31 LybPheGluSerLeuPheArgThrTyraAspLyAspThrThrPheGlnTyrPheYasser 50

```

```

Db 174 AAATTGAGTCCCTCTTAGAGCATGACAGACATCACCTTCAATTTTAAGAGC 233
Oy 51 PheLyAsArgValArgIlaAsnPhSerAsnProLeuSerAlaIaAspAlaArgLeuArg 70
Db 234 TTCAAAACGAGTGAAGATTAACCTTCAGAACCCCTTCCGACGAGATGCGAGCTCCAG 293
Oy 71 LeuHlaLyThrGluPheLeuGlyLyAsGluMeLyLeuTyTyPheAlaGlnThrLeuHla 90
Db 294 CTGCAATAGACGACGACTTCTGAGAAAGAAATGAAGTTATTTTGTCTCAACCTTAAAC 353
Oy 91 IleGlySerSerHlaLeuAlaProProAsnProAspLyGlnPheLeuIleSerProPro 110
Db 354 ATAGAGAGCTCACACCTGCTCCGCAATCCAGACAGACAGTTTCTGATCTCCCTCC 413
Oy 111 AlaserProProValGlyTyrPlyGlnValGluArAlaThrProValIlaAsnTyArAsp 130
Db 414 GCTCTCCGACGAGTGGATGGAACAGTGAAGATGGACCCCACTAATACTATGAT 473
Oy 131 LeuLeuTyraIaIleSerLyLeuGlyProGlyGluLyTyGluLeuHlaIaIaIaThr 150
Db 474 CTCTTATATGCTCATCTCCAACTGGGCGCAGGCAAAAGTATGAATTCACGCGACGACT 533
Oy 151 AspProThrProSerValValIaIaIaValCyAGLuserAspGlnGluangluGluGlu 170
Db 534 GACACCACTCCCGCTGGTGGTCCATGTATGTAGAGTGAATCAAG--AAGAGAGAA 590
Oy 171 GluGluGluMeGluAArgMeLyAsArgProLyProLyLeIleIleGlnThArArArgPro 190
Db 591 GAAGAGAAATGAAAGATGAGAGAGACTTAAGCCAAAATTATTCAGACGAGAGAGCCG 650
Oy 191 GluTyThrProIleHlaIeuser 198
Db 651 GAGTACACGCCGATCCACCTCAGC 674

RESULT 14
B0278576 946 bp mRNA linear EST 07-MAY-2002
LOCUS AGENCOURT_6773220 NIH_MGC_128 Homo sapiens cDNA clone IMAGE:5811711
DEFINITION 5', mRNA sequence.
ACCESSION B0278576
VERSION B0278576.1 GI:20488772
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LUCM2060 row: p column: 16
High quality sequence stop: 548.
Location/Qualifiers
1..946
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5811711"
/clone_lib="NIH_MGC_128"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/notes="Vector: pDNR-LIB; Site:1: SfiI (ggcccatatggcc);
Site:2: SfiI (ggcgccctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA RNAs (bladder

```

- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%. 5' and 3' adaptors were used in cloning as follows:

5'-AAGCAGTGGTATACAGCAGAGTGGCCATTACGCGGGG-3' and 5'-ATTAGAGCCGAGCGGCGGACATG-3' (30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH\_MGC\_126 and NIH\_MGC\_127). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC Library."

BASE COUNT 264 a 243 c 239 g 200 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 3,378-87 Length: 946  
Score: 865.50 Matches: 167  
Percent Similarity: 92.02% Conservative: 6  
Best Local Similarity: 88.83% Mismatches: 14  
Query Match: 82.59% Indels: 1  
DB: 14 Gaps: 1

US-09-782-953-3 (1-198) x BQ278576 (1-946)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
Db 106 AGCTTCCTGATTCCTGTGGCAACAGTGATATCTTCAGCGAAGAGTGAACACAGGGCC 165  
Qy 31 LysPheGluSerLeuPheArgThrTyAspLysAspThrPheGlnTyrPheLysSer 50  
Db 166 AAATTGAGTCCCTCTTAGACGCTATGACAAAGACATCACCTTTTCAGTATTTTAAAGC 225  
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
Db 226 TTCAACAGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 285  
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
Db 286 CTGATAGACTGAGTTCTCGGAAAGAAATGAAGTTATATTTTGTCTGACCTTACAC 345  
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110  
Db 346 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGATTCTGATCTCCCTCCC 405  
Qy 111 AlaSerProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
Db 406 GCCTCTCCGCGAGTGGATGGAACCAAGTGAAGATGCCACCCAGTCATAAACTATGAT 465  
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
Db 466 CTCATTATATGCATCTCCAAAGCTGGGCGGCGGAAAGATGAATGTCAGCGAGGACT 525  
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
Db 526 GACACCACTCCAGCGTGGTCCATGATGTAGAGTGTATCAAGAAAC--GAGGAA 582  
Qy 171 GluGluGluMetGluArgMetLysArgProLysIleIleGlnThrArgArgPro 190  
Db 583 GAAGAGGAATGGAAGAATGACGAGACTTAAGCCAAAATTTATCCCGACCGAAGCGC 642  
Qy 191 GluTyrThrProIleHisLeuSer 198  
Db 643 GAGTACACGCCGATCCACCTCAGC 666

#### RESULT 15

LOCUS AUI31040 718 bp mRNA linear EST 01-AUG-2002  
DEFINITION AUI31040 NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA  
sequence.  
ACCESSION AUI31040

#### VERSION KEYWORDS SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AUI31040.1 GI:10991394

EST.

human.

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (Bases 1 to 718)

Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,

Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

Isogai,T.

HRI human cDNA project

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1..718

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="NT2RP3001895"

/cell\_type="teratocarcinoma"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor

cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 217 a 174 c 169 g 155 t

ORIGIN

Alignment Scores:

Pred. No.: 4,168-87 Length: 718

Score: 863.50 Matches: 166

Percent Similarity: 92.55% Conservative: 8

Best Local Similarity: 88.30% Mismatches: 13

Query Match: 82.40% Indels: 1

DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AUI31040 (1-718)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30

Db 132 AGCTCCTGATTCCTGTGGCAACAGTGATATCTTCAGCGAAGAGTGAACACAGGGCC 191

Qy 31 LysPheGluSerLeuPheArgThrTyAspLysAspThrPheGlnTyrPheLysSer 50

Db 192 AAATTGAGTCCCTCTTAGACGCTATGACAAAGACATCACCTTTTCAGTATTTTAAAGC 251

Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70

Db 252 TTCAACAGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 311

Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90

Db 312 CTGATAAGACTGAGTTCTCGGAAAGAAATGAAGTTATATTTTGTCTGACCTTACAC 371

Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110

Db 372 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGATTCTGATCTCCCTCCC 431

Qy 111 AlaSerProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130

Db 432 GCCTCTCCGCGAGTGGATGGAACCAAGTGAAGATGCCACCCAGTCATAAACTATGAT 491

Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150

Db 492 CTCTTATATGCCATCTCCAAAGCTGGGCGGCGGAAAGATGAATGTCAGCGAGGACT 551





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: December 11, 2002, 11:37:40 ; Search time 13.0219 Seconds  
(without alignments)  
1461.736 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQPSATIACHLDP.....RPKPKLIQTRREYTPIHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:.\*  
1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	355	33.9	239	T34305	hypothetical prote
2	139.5	13.3	315	T49641	nebula related pro
3	95.5	9.1	163	S62444	conserved hypothet
4	94	9.0	318	C84651	hypothetical prote
5	92	8.8	992	T46337	hypothetical prote
6	91.5	8.7	1274	D84485	probable retroelem
7	88	8.4	284	B48516	surfactant protein
8	88	8.4	630	S29796	hypothetical prote
9	88	8.4	721	S29795	hypothetical prote
10	86.5	8.3	822	S30198	Na+/H+-exchanging
11	86	8.2	338	S73902	UDP-glucose 4-epim
12	84.5	8.1	719	S51739	transcription repr
13	83	7.9	561	JC2436	5'-nucleotidase (E
14	83	7.9	1173	T31421	C-terminal domain-
15	82	7.8	697	S96752	hypothetical prote
16	82	7.8	1477	T13797	tumor suppressor pr
17	81.5	7.8	528	C69323	conserved hypothet
18	81	7.7	151	T10768	latex allergen Hev
19	80.5	7.7	280	C64471	hypothetical prote
20	80.5	7.7	360	S12850	protein TPX - Ther
21	80	7.6	681	T17342	hypothetical prote
22	79.5	7.6	797	A70453	glutamate ammonia
23	79.5	7.6	904	I38757	homolog of Drosoph
24	79.5	7.6	926	I38756	homolog of Drosoph
25	79.5	7.6	1020	B86414	hypothetical prote
26	79.5	7.6	1136	T26953	hypothetical prote
27	79.5	7.6	1146	C83304	hypothetical prote
28	79.5	7.6	1557	T23132	hypothetical prote
29	79	7.5	684	F85075	hypothetical prote

protein-tyrosine k  
spindle assembly c  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
zinc finger protei  
actin - medicina  
acetyl-CoA carboxy  
involucrin - mouse  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
dnak-type molecula  
AF4 protein - mous  
hypothetical prote

30 79 7.5 1154 2 A39577  
31 79 7.5 1216 2 T09224  
32 79 7.5 13055 2 T16580  
33 78.5 7.5 211 2 S37792  
34 78.5 7.5 290 2 T42526  
35 78.5 7.5 1706 2 I84499  
36 78.5 7.5 1880 2 T18531  
37 78 7.4 454 2 AE1244  
38 78 7.4 467 1 A49377  
39 78 7.4 441 2 A8151  
40 77.5 7.4 455 2 B86427  
41 77.5 7.4 621 2 T15929  
42 77.5 7.4 627 2 A41609  
43 77.5 7.4 1211 2 T42230  
44 77 7.3 340 2 B84771  
45

RESULT 1  
T34305  
hypothetical protein F54E7.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34305  
R:Bentley, D.  
submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of C. elegans cosmid F54E7.  
A:Reference number: Z21502  
A:Accession: T34305  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-239 <BEN>  
A:Cross-references: EMBL:U00067; PIDN:AAC7519.1; GSPDB:GN00021; CESP:F54E7.7  
A:Experimental source: strain Bristol N2; clone F54E7  
C:Genetics:  
A:Gene: CESP:F54E7.7  
A:Map position: 3  
A:Introns: 40/3; 78/2; 165/2; 205/1

Query Match 33.9%; Score 355; DB 2; Length 239;  
Best Local Similarity 41.0%; Pred. No. 2.3e-24;  
Matches 75; Conservative 36; Mismatches 68; Indels 4; Gaps 3;

QY 8 DLPSATIAIHLDPVFDGLCRKAFESLFRITYDKDTTFQYFKSKRVINFSNPLGSAADA 67  
Db 27 DLPNAIIVTQVPEDYFDNKNQKDNKANSFLFTQIEKDIHFDLRSFRRVRVIFSSPENATAA 86  
QY 68 RLRLHKTEFLGKEMKLYFAQTUHI--GSHLAPPNPDQFLISPPASPPVGVKQVEDATP 125  
Db 87 KLIVQFSGFKGHELKAFQAQRIYMSANSQMLSPFPLEKQFLISPPCSPVGVGEQTKMPP 146  
QY 126 VI-NYDLLVAISKLGKGEKVELHAATDPTSPVVHVCSDDQNEEEEEEMERMKRPKPKI 184  
Db 147 WCNFDMARLASFAIDEXIEVHNGDELTPAIVHPCETPID-VPSAIEMPRTTPRSSPS 205  
QY 185 IQT 187  
Db 206 LQS 208

RESULT 2  
T49641  
nebula related protein [imported] - Neurospora crassa  
N:Alternate names: protein B5022.230  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49641  
R:Schulte, U.; Align, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022

A:Accession: T49641  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-315 <SCH>  
 A:Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.230  
 A:Experimental source: BAC clone B5022; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B5022.230  
 A:Map position: 6

Query Match 13.3%; Score 139.5; DB 2; Length 315;  
 Best Local Similarity 40.2%; Pred. No. 6.4e-05;  
 Matches 39; Conservative 13; Mismatches 38; Indels 7; Gaps 3;

QY 49 KSPKRVINPSPUSADARLRHKTFLGKMKLYPAQTLHGSS----HLAPPNDKQ 104  
 DB 101 KSPRIITVFEDEQAALAVRSVMDGEALIGRCRVFGQPRPIDVSADHMLPDKGL 160  
 QY 105 FLISPPSPVGMKQVEDA--TPVINYDLIAISK 138  
 DB 161 FLISPPSPVGMKQVEDA--TPVINYDLIAISK 197

# RESULT 3

conserved hypothetical protein SPAC13G6.15C - fission yeast (Schizosaccharomyces pombe)  
 N:Alternate names: hypothetical protein SPAC24B11.04C  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #ext\_change 10-Dec-1999  
 C:Accession: S62444; S62549; T37650; T38331  
 R:Odell, C.; Bowman, S.  
 Submitted to the EMBL Data Library, October 1995

A:Reference number: S62430  
 A:Accession: S62444  
 A:Molecule type: DNA  
 A:Residues: 1-163 <ODE>  
 A:Cross-references: EMBL:Z54308; NID:g1008985; PIDN:CAA91108.1; PID:g1009000  
 R:Odell, C.; Churcher, C.M.  
 Submitted to the EMBL Data Library, November 1995  
 A:Reference number: S62546  
 A:Accession: S62549  
 A:Molecule type: DNA

A:Residues: 1-163 <OD2>  
 A:Cross-references: EMBL:Z67757; NID:g1061288; PIDN:CAA91769.1; PID:g1061292  
 R:Odell, C.; Bowman, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.  
 Submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z21734  
 A:Accession: T37650  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T37650  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-163 <OD3>  
 A:Cross-references: EMBL:Z54308; PIDN:CAA91108.1; GSPDB:GN00066; SPDB:SPAC13G6.15C  
 R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 Submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z21786  
 A:Accession: T38331  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-163 <ODA>  
 A:Cross-references: EMBL:Z67757; PIDN:CAA91769.1; GSPDB:GN00066; SPDB:SPAC24B11.04C  
 A:Experimental source: strain 972h-; cosmid c24B11  
 C:Genetics:  
 A:Gene: SPAC24B11.04C; SPAC13G6.15C; SPAC24B11.04C  
 A:Map position: 1L  
 C:Superfamily: fission yeast hypothetical protein SPAC13G6.15C

Query Match 9.1%; Score 95.5; DB 2; Length 163;  
 Best Local Similarity 32.1%; Pred. No. 0.25;  
 Matches 25; Conservative 17; Mismatches 25; Indels 11; Gaps 4;  
 QY 96 LAPRPNDKQFLISPPSPVGMKQVEDATP---VINIDLIAISKLGKGYELHAATDP 152

DB 85 LGVRFKFKMILSPSPSPVGMKQVEDATPSPNSOHLAHDIOQLDELGNALLND-HSA--- 140  
 QY 153 TPSVVVHVCESDQENEE 170  
 DB 141 GPQIVL-----SEHNNTKE 154

# RESULT 4

hypothetical protein At2g25670 (imported) - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #ext\_change 02-Feb-2001  
 C:Accession: C84651  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 Eaux, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999  
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: C84651  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <STO>  
 A:Cross-references: GB:A802093; NID:g4874305; PIDN:AMD1367.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g25670  
 A:Map position: 2

Query Match 9.0%; Score 94; DB 2; Length 318;  
 Best Local Similarity 25.2%; Pred. No. 0.77;  
 Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

QY 54 VRINSPNLSAADARLRHKTFLGKMKLYPAQTLHGSHLAPP--NDKQFLISPPA 111  
 DB 13 ITTOSTNLPALDTRKKKKSDKXGSK-----GSSKSRBEKEPEPVYVAPTP 62

QY 112 SPVGMKQVEDATPVINYDLIAISKLGKGYELHAATDPSPVVVHVCESDQ----- 166  
 DB 63 LKVKSMADIDEDDDDDYVATTPAQSGWSTLPSH-----TDSKVHVESESEEDILD 117

QY 167 -----NEEBEEMERKRPKPKITQTRPE 191  
 DB 118 EGDVDEBEQETEVQVHPEPEV--KKAPE 146

# RESULT 5

hypothetical protein DXFPZ43402413.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #ext\_change 04-Feb-2000  
 C:Accession: T46337  
 R:Koehner, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 Submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23037  
 A:Accession: T46337  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-992 <AAA>  
 A:Cross-references: EMBL:AL137265  
 A:Experimental source: adult testis; clone DXFPZ43402413  
 C:Genetics:  
 A:Note: DXFPZ43402413.1

Query Match 8.8%; Score 92; DB 2; Length 992;  
 Best Local Similarity 22.7%; Pred. No. 4.7;  
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;

QY 96 LAPRPNDKQFLISPPSPVGMKQVEDATPVINYDLIAISKLGKGYELHAATDPSPV----- 136  
 DB 245 LSPFLPHERASPPSLATEEPQEGEGQEWKAE-----LGEDSAASLSQDLSOR 300

QY 137 -----KLGRG-----EKYELHAATDPSPV----- 156

Db 301 EQAPSPAAACEKGEKSOAEELGPOEAEDEPEEKVAVSPPTPPSPVSRSTEPVAPPEQ 360  
 QY 157 -----VVHVCSQDQNEEEEMERMKPKPKIOTRRPEYTPIH 196  
 Db 361 LSEAALKAMEEAAVQLEODQRHLLESKO-EKMQQLREKLQCEEEIEILRLH 411  
 RESULT 6  
 Db4485  
 probable retroelement pol polyprotein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84485  
 R.; Sin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;  
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84485  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1274 <STO>  
 A:Cross-references: GB:AE002093; NID:94309763; PIDN:AAI:5532.1; GSPDB:GN00139  
 C:Genetics:  
 A:Genome: chloroplast  
 A:Map position: 2

Query Match 8.7%; Score 91.5; DB 2; Length 1274;  
 Best Local Similarity 23.7%; Pred. No. 7.2;  
 Matches 51; Conservative 30; Mismatches 83; Indels 51; Gaps 9;  
 QY 24 VDGLCRAPFSLFRTYDKDTTFQYFKSKRVRINFNSPLSAADAR-LRLHKTFFLGKE-- 80  
 Db 673 IESRVEAKFEGRTGSIENDV-----KQKEQLKAIADKSSSVYRDWYLAKTQPTQODN 727  
 QY 81 MKLYFAOTLHI-----GSHLAPNPD-----KQFLISPPA-----SPPVGWK 118  
 Db 728 PKVQTQTPDPVPKPTNQATSPSPSKQADVKKETLNELIQPKPGRGRKPSQOPK 787  
 QY 119 QVEDATPVINDLLYALSKLGGKGYELHA-----ADTPSPVVHVHVCSDQNEEEEEE 173  
 Db 788 KVSPTMPKX-----TIKKLQPSQAEKAEEDSDVDVDTKVVSEYNMLPESDEDEEE 841  
 QY 174 MERMKRPKPKIOTRR-----PEYTPIHLS 198  
 Db 842 TERLKSVEIRLKTVKLAPDGRSLINPEAMPHTS 876  
 RESULT 7  
 B48516  
 surfactant protein A receptor - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 07-Apr-1994  
 C:Accession: B48516  
 R.; Strayer, D.S.; Yang, S.; Jerng, H.H.  
 J. Biol. Chem. 268, 18679-18684, 1993  
 A:Title: Surfactant protein A-binding proteins. Characterization and structures.  
 A:Reference number: A48516; MUID:93366778; PMID:8360162  
 A:Accession: B48516  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-284 <STR>  
 A:Cross-references: GB:L10124

Query Match 8.4%; Score 88; DB 2; Length 284;  
 Best Local Similarity 24.1%; Pred. No. 2.3;  
 Matches 40; Conservative 24; Mismatches 62; Indels 40; Gaps 7;  
 QY 27 LCRAKFSLEFTYDKDTTFQYFKSKRVRINFNSPLSAADARLRLHKTFFLGKEMKLYEA 86  
 Db 3 MCYGMGSLFRSGSRQTLFA-----SQMRYAD-----LYAASFIN-----LLYYP 43

QY 87 QTLHGSSHLAPPNP-----DKQFLISPPAS-----PPVGWKQVEDATPVINDLLYAI 135  
 Db 44 FSYLFAARHVLMPHSTVEHTHVDINEMESPLATNRTSVDFKD-----TDYKRHQLTRSI 99  
 QY 136 SKLGGKGYELHAATDPTSPVVHVHVCSDQNEEEEMERMKRPK 181  
 Db 100 SEIKPPNLPPL-----APQEITHCHDEDDDEEEEEECECGRPK 139  
 RESULT 8  
 S29796  
 hypothetical protein 2280 - evening primrose chloroplast (fragment)  
 C:Species: chloroplast Oenothera villaricae (evening primrose)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
 C:Accession: S29796; S19986  
 R.; Nimzyk, R.; Schoendorf, T.; Hachtel, W.  
 Curr. Genet. 23, 265-270, 1993  
 A:Title: In-frame length mutations associated with short tandem repeats are located in  
 A:Reference number: S29795; MUID:93169690; PMID:8435856  
 A:Accession: S29796  
 A:Molecule type: DNA  
 A:Residues: 1-630 <NIM>  
 A:Cross-references: EMBL:X64615; NID:g14331; PID:g14332  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast  
 Query Match 8.4%; Score 88; DB 2; Length 630;  
 Best Local Similarity 26.2%; Pred. No. 6.2;  
 Matches 45; Conservative 23; Mismatches 48; Indels 56; Gaps 10;  
 QY 16 CHLDP-RVFPV-DGLCRAPFSLFRTYDKDTTFQYFKSKRVRINFNSPLSAADARLRLHK 73  
 Db 81 CSIDPISIVIKNLCAGDSLSLYKY-----FELGTSMKKLT----- 118  
 QY 74 TFFLGKEMKLYFAQTLHIGS---SHLAPPNPKQFLISPPSPVVGWKQVEDATPVINDY 130  
 Db 119 -----LLYLLTCSAGSIAQDLLSPGPDEQNLT-----SYGLVEN-----DSD 157  
 QY 131 LLYAISKLGKGYELHA-----TDTPSPVVHVHVCSDQNEEE-----EEME 175  
 Db 158 LVHGLSDIVHG-LLELEGALVGSPTTEBEVEGTEBEVEGTEBEVE 208  
 RESULT 9  
 S29795  
 hypothetical protein 2280 - evening primrose (Oenothera picensis subsp. picensis) chlor  
 C:Species: chloroplast Oenothera picensis subsp. picensis (evening primrose)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
 C:Accession: S29795; S19983  
 R.; Nimzyk, R.; Schoendorf, T.; Hachtel, W.  
 Curr. Genet. 23, 265-270, 1993  
 A:Title: In-frame length mutations associated with short tandem repeats are located in  
 A:Reference number: S29795; MUID:93169690; PMID:8435856  
 A:Accession: S29795  
 A:Molecule type: DNA  
 A:Residues: 1-721 <NIM>  
 A:Cross-references: EMBL:X64616; NID:g14334; PID:g14335  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast  
 Query Match 8.4%; Score 88; DB 2; Length 721;  
 Best Local Similarity 26.2%; Pred. No. 7.3;  
 Matches 45; Conservative 23; Mismatches 48; Indels 56; Gaps 10;  
 QY 16 CHLDP-RVFPV-DGLCRAPFSLFRTYDKDTTFQYFKSKRVRINFNSPLSAADARLRLHK 73  
 Db 81 CSIDPISIVIKNLCAGDSLSLYKY-----FELGTSMKKLT----- 118  
 QY 74 TFFLGKEMKLYFAQTLHIGS---SHLAPPNPKQFLISPPSPVVGWKQVEDATPVINDY 130  
 Db 119 -----LLYLLTCSAGSIAQDLLSPGPDEQNLT-----SYGLVEN-----DSD 157

Oy 131 LLYAISKLGPEKYEELHAA--TDPTPSVVHVHVCSDQENEER-----EEEME 175  
 Db 158 LVHGLSDIVHG-LLELEGALVGSPTVEEVECTEEVECTEEVECTEEVEE 208

# RESULT 10

S30198  
 Na+/H+-exchanging protein - Chinese hamster  
 N.Alternate names: Na+/H+ antiporter; Na+/H+ exchanger  
 C.Species: Cricetus griseus (Chinese hamster)  
 C.Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 05-Nov-1999  
 C.Accession: S30198; #sequence\_revision 30-Sep-1993 #text\_change 05-Nov-1999  
 R.Coumilion, L.; Pouyssegur, J.  
 Biochim. Biophys. Acta 1172, 343-345, 1993  
 A.Title: Nucleotide sequence of the Chinese hamster Na(+)/H(+) exchanger NHE1.  
 A.Reference number: S30198; MUID:93192332; PMID:8383540  
 A.Accession: S30198  
 A.Molecule type: mRNA  
 A.Residues: 1-822 <CON>  
 A.Cross-references: EMBL:X68970; NID:949472; PID:CAA8771.1; PID:949473  
 C.Genetics: NHE1  
 C.Keywords: glycoprotein; ion transport; membrane protein  
 F:76,374,414/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 8.3%; Score 86.5; DB 2; Length 822;  
 Best Local Similarity 23.7%; Pred. No. 12;  
 Matches 40; Conservative 26; Mismatches 64; Indels 39; Gaps 8;

Oy 44 TQYFSPKRVIRINFSNPLSAA--DKRLRLKTEFLGKEMKLYFAQTLLH----- 90  
 Db 646 TRKRLSYNNHTL-VADPYBEANNQMLRRQKARQLEQKSNLYTPAHKLDSPTRRAR 704  
 Oy 91 TGSNHLA-PNPDKQFLSPSPAPVGMKQVEDATPVINVDLLVAISKLGPEKYEELHAA 149  
 Db 705 TGSDFLALYPRKADLPVITIDPASP-----QSPESVDLVNELKAKV-----LGVN 749  
 Oy 150 TDPTPSVVHVHVCSDQENEEREMERMKRPKPKI---IQTRRPETYP 194  
 Db 750 RDPT-----RLTRGEDEDEDEGVIIMRKKEPSSPGTDTFTAPWYSP 793

# RESULT 11

S73902  
 UDP-glucose 4-epimerase gale - Mycoplasma pneumoniae (strain ATCC 29342)  
 N.Alternate names: hypothetical protein A65\_orf338  
 C.Species: Mycoplasma pneumoniae  
 A.Variety: ATCC 29342  
 C.Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
 C.Accession: S73902  
 R.Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A.Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A.Reference number: S73327; MUID:97105885; PMID:894633  
 A.Accession: S73902  
 A.Structus: preliminary; nucleic acid sequence not shown; translation not shown  
 A.Molecule type: DNA  
 A.Residues: 1-338 <HIN>  
 A.Cross-references: EMBL:AF000056; GB:U00089; NID:91674263; PID:AB96224.1; PID:9167427  
 A.Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C.Genetics:  
 A.Gene: gale  
 A.Genetic code: SGC3  
 C.Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 F:8-338/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 8.2%; Score 86; DB 2; Length 338;  
 Best Local Similarity 27.9%; Pred. No. 4.3;  
 Matches 34; Conservative 19; Mismatches 59; Indels 10; Gaps 3;  
 Oy 4 VDLQDLPSATIGHDPRV---FVGLCRAKFESLFTYDQDTFOYKSPKRVIRINS 59  
 Db 158 LVHGLSDIVHG-LLELEGALVGSPTVEEVECTEEVECTEEVECTEEVEE 208

Db 34 VSVIDNHTSLALLPQVNVHFNLLDRAQLFTTIAQINPDVVF-HFAAKTSVKESTE 92  
 Oy 60 NPLSADARL-----RLKTEFLGKEMKLYFAQTLLHISGSLAPNPDKQFLSPAP 114  
 Db 93 QPLTFPDHVLVGTLLNLHLAKELQKIQLPFSSTAVFGSASTYPIENLVLEETLASNP 152

Oy 115 VG 116  
 Db 153 YG 154

# RESULT 12

S51739  
 transcription repressor AEBP1 - mouse  
 C.Species: Mus musculus (house mouse)  
 C.Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 08-Oct-1999  
 C.Accession: S60227; S51739  
 R.He, G.P.; Mulise, A.; Li, A.W.; Ro, H.S.  
 Nature 378, 92-96, 1995  
 A.Title: A eukaryotic transcriptional repressor with carboxypeptidase activity.  
 A.Reference number: S60227; MUID:96061010; PMID:7477239  
 A.Accession: S60227  
 A.Structus: preliminary  
 A.Molecule type: mRNA  
 A.Residues: 1-719 <HE2>  
 A.Cross-references: EMBL:X80478; NID:9607131; PID:CAA56648.1; PID:9607132

Query Match 8.1%; Score 84.5; DB 2; Length 719;  
 Best Local Similarity 21.8%; Pred. No. 15;  
 Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

Oy 54 VRNFSNPLSADARLRLKTEF--LGKEMKLYFAQTLLHISGSLAPNPDKQFLSP 110  
 Db 585 LRVDPSRPMTPQQRNMQRRLQYRLRMREQRL--RLNSTAPATSPPT---ALMPP 637  
 Oy 111 ASP-----PVGWK--QVEDATPVINVDLLVAISKLGPEKYEELHAAIDPT 153  
 Db 638 PSPPTAATLRPMEVLPPTTAGMSESETEYTEVVT-----EPETEYGD-- 681  
 Oy 154 PSVVHVHVCSDQENEEREMERMKRPKPKI---IQTRRPETYP 194  
 Db 682 ----LVEVEIEEEREEEND 699

# RESULT 13

JC2436  
 5'-nucleotidase (BC 3.1.3.5), cytosolic - human  
 N.Alternate names: cytosolic purine 5'-nucleotidase  
 C.Species: Homo sapiens (man)  
 C.Date: 15-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 21-Jul-2000  
 C.Accession: JC2436  
 R.Oka, J.; Matsumoto, A.; Hosokawa, Y.; Inoue, S.  
 Biochem. Biophys. Res. Commun. 205, 917-922, 1994  
 A.Title: Molecular cloning of human cytosolic purine 5'-nucleotidase.  
 A.Reference number: JC2436; MUID:95091838; PMID:7999131  
 A.Accession: JC2436  
 A.Molecule type: mRNA  
 A.Residues: 1-561 <OKA>  
 A.Cross-references: DDBJ:D38524; NID:9633070; PID:BA07529.1; PID:9633071  
 C.Comment: This enzyme is a soluble nucleotidase. It hydrolyzes IMP and other purine nuc.  
 C.Comment: This enzyme has a critical role in the maintenance of a constant composition,  
 C.Genetics:  
 A.Gene: GDB:NTSCP  
 A.Cross-references: GDB:569367  
 C.Keywords: cytosol; phosphoric monoester hydrolase

Query Match 7.9%; Score 83; DB 2; Length 561;  
 Best Local Similarity 23.3%; Pred. No. 15;  
 Matches 47; Conservative 28; Mismatches 69; Indels 58; Gaps 10;  
 Oy 2 EEVDLQDLPSATIGHDPRV---PRFVD-GLCRAKFEELFTYDQDT 44  
 Db 390 BEGLSDIFLAELKXHLSSNSNERPDISSIORRIKVTHTMDWCYGMGSLFRSGRQTL 449



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1655.79 Seconds  
(without alignments)  
3480.130 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQDLPSATIACHLDP.....RPKXLIQTRRPEYTPHLS 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh  
-Q/cgn2\_1/USPTO.spool/US09782953/runat 11122002 114429 17340/app query.fasta\_1.1173  
-DB=GenEmbl -QFMT=fastap -SUPFIX=rge -MINMATCH=0\_1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bite -START=1 -END=1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 -CGCN 1 2844 @runat 11122002 114429 17340 -NCPU=6 -ICPU=3  
-NO XUPFY -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb\_ba :  
2: gb\_htg :  
3: gb\_in :  
4: gb\_ov :  
5: gb\_ov :  
6: gb\_pat :  
7: gb\_ph :  
8: gb\_pl :  
9: gb\_pr :  
10: gb\_ro :  
11: gb\_sca :  
12: gb\_sy :  
13: gb\_un :  
14: gb\_vi :  
15: em\_ba :  
16: em\_fun :  
17: em\_hum :  
18: em\_in :  
19: em\_mu :  
20: em\_on :  
21: em\_ox :  
22: em\_ov :  
23: em\_pat :  
24: em\_ph :  
25: em\_pi :  
26: em\_ro :  
27: em\_sts :  
28: em\_un :

29: em\_vi :  
30: em\_htg\_hum :  
31: em\_htg\_inv :  
32: em\_htg\_other :  
33: em\_htg\_mus :  
34: em\_htg\_pln :  
35: em\_htg\_rod :  
36: em\_htg\_mam :  
37: em\_htg\_vrt :  
38: em\_sy :  
39: em\_hcgo\_hum :  
40: em\_hcgo\_mus :  
41: em\_hcgo\_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	597	6	AX365312 Sequence
2	1048	100.0	597	10	AF237789 Mus muscu
3	1040	99.2	621	10	AF282255 Mus muscu
4	1040	99.2	2125	10	AF260717 Mus muscu
5	1002.5	95.7	2212	6	AX365324 Sequence
6	1002.5	95.7	2227	9	HSU85266
7	1002.5	95.7	2289	9	BC002864 Homo sapi
8	1002.5	95.7	2407	9	AK092184 Homo sapi
9	920.5	87.8	597	6	AX365315 Sequence
10	920.5	87.8	597	10	AF237790 Mus muscu
11	911	86.9	2141	10	AF263239 Mus muscu
12	911	86.9	2198	10	BC033551 Mus muscu
13	890	84.9	2224	10	AF263240 Mus muscu
14	887.5	84.7	626	10	AB075973 Rattus no
15	884.5	84.4	2216	10	CGU60263 Cricetulus
16	872.5	83.3	2331	6	AX365321 Sequence
17	872.5	83.3	2346	9	HSU85267 Homo sapien
18	859.5	82.0	2348	6	AX281651 Sequence
19	847.5	80.9	2173	6	AX410694 Sequence
20	847.5	80.9	2173	9	HSU28833 Homo sapien
21	847.5	80.9	2284	9	HSU85265 Homo sapien
22	847	80.8	2679	9	AK055845 Homo sapi
23	844.5	80.6	2174	6	AR034241 Sequence
24	839.5	80.1	798	9	AF400429 Homo sapi
25	749.5	71.5	599	6	AX365311 Sequence
26	709	67.7	562	9	HSU53821 Homo sapien
27	627.5	59.9	3295	10	AB061525 Mus muscu
28	620.5	59.2	934	6	AX074351 Sequence
29	620.5	59.2	3159	6	AX420425 Sequence
30	620.5	59.2	3261	9	AK090990 Homo sapi
31	613.5	58.5	594	6	AX365318 Sequence
32	613.5	58.5	594	10	AF237791 Mus muscu
33	613.5	58.5	659	10	AF237887 Mus muscu
34	608.5	58.1	3227	10	AB061524 Mus muscu
35	607.5	58.0	3240	9	AY034085 Homo sapi
36	607.5	58.0	3253	9	AY034086 Homo sapi
37	597	57.0	776	10	AF237888 Mus muscu
38	596.5	56.9	3184	6	AX329596 Sequence
39	596.5	56.9	3184	6	AX333015 Sequence
40	596.5	56.9	3184	6	AX365327 Sequence
41	596.5	56.9	3184	6	AX420436 Sequence
42	596.5	56.9	3184	9	HUM2AKI4
43	596	56.9	828	6	AX365330 Sequence
44	596	56.9	828	6	AX420437 Sequence
45	596	56.9	828	9	AF176116 Homo sapi

# ALIGNMENTS

RESULT 1

AX365312	LOCUS	AX365312	597 bp	DNA	linear	PAT 15-FEB-2002
AX365312	DEFINITION	Sequence 2 from Patent WO0204491.				
AX365312	ACCESSION	AX365312				
AX365312.1	VERSION	AX365312.1	GI:18697043			
KEYWORDS						
SOURCE						
ORGANISM		house mouse.				
		Mus musculus				
		Euarchonta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE						
AUTHORS		1 Williams, S.R. and Rothermel, B.				
TITLE		Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)				
JOURNAL		Patent: WO 0204491-A 2 17-JAN-2002;				
		Board of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rothermel, Beverly (US)				
FEATURES						
Source		Location/Qualifiers				
		1..597				
		/organism="Mus musculus"				
		/db_xref="taxon:10090"				
		1..597				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAD23808.1"				
		/db_xref="GI:18697044"				
		/translation="MEVDLDQDPEATACHLDPEVAVDGLCRAKFEELFRTYKDTT				
		/translaltion="FOYFKSFKRVRIENFNPISADARLRHLKTEBLGKEMLYFAQTHISSHLAPNP				
		KQFLSPSPSPVGVKQVEDATPVINYLVAISLGGEXELMAADPPSPVVVHH				
		CESDENESEEEEMERMKRPKRIQTRPEYTPHLS"				
BASE COUNT		160 a 170 c 151 g 116 t				
ORIGIN						
Alignment Scores:						
Pred. No.:		5,066-96	Length:	597		
Score:		1048.00	Matches:	198		
Percent Similarity:		100.00%	Conservative:	0		
Best Local Similarity:		100.00%	Mismatches:	0		
Query Match:		100.00%	Indels:	0		
DB:		6	Gaps:	0		
US-09-782-953-3 (1-198) x AX365312 (1-597)						
QY	1	MctGluGluValAspLeuGlnAaPleuGlnAaPleuProSerAlaThrIleAlaCysHisLeuAaPPro	20			
Db	1	ATGGAGGAGGAGTGGATTGTCAGACCTGGCCAGCCACCATCGCTCCACCTGGACCCG	60			
QY	21	ArgValAPheValAaPLeuGlyLeuGlyCysArgAlaAlaPheGlnSerLeuPheArgThrTYzAaP	40			
Db	61	CGCGTGTTCGTGGACGGCGCTGTCGGGGCCAAATTGGAAATCCCTCTTCAGAACATATGAC	120			
QY	41	LyAsaPThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAaP	60			
Db	121	AAGGACACCACTTCCTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATTAACCTCAGCAAC	180			
QY	61	ProLeuSerAlaIleAaPAlaArgLeuArgLeuHisIbysThrGlnPheLeuGlyLySgln	80			
Db	181	CCCTTATCTGCACCCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGAGAGAA	240			
QY	81	MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerThrIleuAlaPProProAaP	100			
Db	241	ATGAAATTGTAATTTTGTCTCAGACTTTTACATATGAGAAATTCACACCTGGCTCCGCCAAT	300			
QY	101	ProAspLysGlnPheLeuLieserProPAlaSerProProValGlyTPLYSGlnVal	120			
Db	301	CCCGACCAACAGTTCCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGAAACAAGTA	360			
QY	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140			
Db	361	GAGAGTCCCAACCCCGCATAAATTAGATCTTTATATCCATCTCCAAAGCTGGGGCCA	420			
QY	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValValHisVal	160			

Db	421	GGAAGGAAATGATGACCTGGTCCCGGCAATTTGAATCCTCTTCAGAACATATGACC	120
Oy	161	CysGUSerAapGInGLuaSngLUglUglUglUglUmeGjwaXmEcLYsaRgPro	180
Db	481	TGTGAGATGACCAAGAGATGAGAGAGAGAGAGATGAGAGAAATGAGAAGACCC	540
Oy	181	LysProlYseIleIleGlnThrArgArgProGUlyrThrrProIleHalaLeuSer	198
Db	541	AAGCCCAAATCATCTCACACACGAGACCGGAGTAGACACCGATCCACTTACGC	594
RESULT 2			
LOCUS	AF237789	597 bp	mRNA linear ROD 12-APR-2000
DEFINITION	Mus musculus myocyte-enriched calcineurin interactin protein 1 splice variant 1 mRNA, complete cds.		
ACCESSION	AF237789		
VERSION	AF237789.1	GI:7542525	
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Rochermel,B.A., Vega,R.B., Yang,J., Wu,H., BaaseJ-Duby,R.S. and Williams,R.S.		
TITLE	A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling J. Biol. Chem. (2000) In press		
JOURNAL	2 (bases 1 to 597)		
REFERENCE	Rochermel,B.A., Vega,R.B., Yang,J., Wu,H., BaaseJ-Duby,R.S. and Williams,R.S.		
AUTHORS	Direct Submission		
TITLE	Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
FEATURES			
Source	Location/Qualifiers		
	1..597		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="16"		
CDS	1..597		
	/note="similar to Homo sapiens DSCR1, MCP1, similar to EST sequence of Genbank Accession Number AA734360"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin interactin protein 1 splice variant 1"		
	/protein_id="AAF63485.1"		
	/db_xref="gi:7542526"		
	/translation="MEEVDLDDPSATIACHLDPRVFDGCRAKFESLPFYDQDTFOYGFSPRRVRIRNFNSNLSADARLRKTEFLGEMKLVTFAQTHTICSSHLAPNPVKOFLSPSPASPVPGKQVEDATPAVINYLVAISLGAEKYELHAADPTSPVVHCESDENEEEEEMERMRPKKIIOTRPEREYPPIHS"		
BASE COUNT	160 a 170 c 151 g 116 t		
ORIGIN			
Alignment Scores:			
Pred. No. :	5,06e-96	Length:	597
Score:	1048.00	Matches:	198
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-782-953-3 (1-198) x AF237789 (1-597)			
Oy	1	MetGlUGluValAspLeuGlnAspLeuPsoSerAlaThrIleAlaCysHisLeuAapPro	20
Db	1	ATGAGAGAGAGGTGATTGCGAGACCTGCCGAGGCCACCAATCGCCTCCACACCGACCCG	60
Oy	21	ArgValPheValAspGlyLeuCySaTgaIalALyPheGluSerLeuPheArgThryrAap	40
Db	61	CGCGTGTCGTGAGACGGCCCTGTCGCCGGCCAATTTGAATCCTCTTCAGAACATATGAC	120





Guinera, J., Arbones, M., Florez, J., Soriano, E., Estivill, X. and Alcantara, S.  
 Descl, a novel endogenous inhibitor of calcineurin signaling, is expressed in the primitive ventricle of the heart and during neurogenesis  
 Mech. Dev. 101 (1-2), 289-292 (2001)

JOURNAL MEDLINE 21152920  
 PUBLISHED 11231093  
 REFERENCE 2 (bases 1 to 2125)  
 AUTHORS Fuentes, J., Pritchard, M., Pucharcas, C. and Estivill, X.  
 TITLE Down syndrome candidate region 1 (Descl), one of three alternatively spliced exon 1  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 2125)  
 AUTHORS Fuentes, J., Pritchard, M., Pucharcas, C. and Estivill, X.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-APR-2000) Centre de Genetica Medica i Molecular, IRO, Avia, Castelldefels Km. 2,7, L'Hospital de Llobregat, Barcelona 08907, Spain

FEATURES  
 source Location/Qualifiers  
 1..2125  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 1..2125  
 /gene="Descl"  
 15..611  
 /gene="Descl"  
 /note="calcineurin inhibitor"  
 /codon\_start=1  
 /product="Down syndrome candidate region 1"  
 /protein\_id="AA07043.1"  
 /db\_xref="GI:7839597"  
 /translation="MEVDLDLPDPSATTACHLDPRVVDGLCRAKFSLFRTYDKDTT  
 FOYKSPKRVINFSNPLSADARLRLHKEFLGKMKLYFAOTLHGSSHLAPNPD  
 KQFLSPSPVGMKQVEDATPVINDLVLSLGGKEKELMAADTTTPSVVHY  
 CESDQKEEBEEMERMRPKKIIQTRPPTPIHLS"

BASE COUNT 487 a 536 c 563 g 536 t 3 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,47e-94 Length: 2125  
 Score: 1040.00 Matches: 197  
 Percent Similarity: 99.49% Conservative: 0  
 Best Local Similarity: 99.49% Mismatches: 1  
 Query Match: 99.24% Indels: 0  
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x AF260717 (1-2125)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHisLeuAspPro 20  
 DB 15 ATGAGAGAGGTGATCTGACGACCTGCCAGGCGCACCATGCTGCACCTGGACCCG 74  
 OY 21 ArgValAlpheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 DB 75 CGCGTGTGTGTGACGCGCTGTGCGGGCCAAATTGAACTCCCTTCGGAACATATGAC 134  
 OY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheserAsn 60  
 DB 135 AAGGACACCACTTCCAGATTTTAAAGACTTCAAAAGTGTCCGGAATAACTTCAGCAAC 194  
 OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 DB 195 CCTTATCTGACGACCTGACGCTGCGGCTGCACAAAGCCAGAGTCTCTGGGAAAGAA 254  
 OY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysLysSerSerHisLeuAlaProProAsn 100  
 DB 255 ATGAAGTGTATTATTGCTACACTTACATAGAAAGTTACACCTGCTCGGCCCAAT 314  
 OY 101 ProAspLysGlnPheLeuLysSerProAlaSerProProValGlyTrrPheGlnVal 120  
 DB 315 CCGGACAAACATTTCTCTCTCCCTCCGCGCTCTCTCCCTGCGGTGGTGGAAACAAAGA 374

OY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
 DB 375 GAAGATGCCACCCCGCTGATTAATTACGATCTTTATATGTCATCTCCAGCTGGGGCCA 434  
 OY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
 DB 435 GGAGAGAGTATGAAGCTGATGACAGCAGACACACTCCAGTGTGTGCTCCACGCG 494  
 OY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
 DB 495 TGTGAGTACGACCAAGATATAGAGAGAGAGAGAGAGAGATGACAGATGAAGAGACC 554  
 OY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
 DB 555 AAGCCCAAAATCATTCACAGACGAGACCGGAGTACACCATCCTCCTCAGC 608

RESULT 5  
 AX365324  
 LOCUS AX365324 2212 bp DNA linear PAT 15-FEB-2002  
 DEFINITION Sequence 14 from Patent WO0204491.  
 ACCSSION AX365324  
 VERSION AX365324.1 GI:18697051  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 Williams, S.R. and Rothermel, B.  
 Methods and compositions relating to muscle selective calcineurin  
 interacting protein (mcip)  
 Patent: WO 0204491-A 14 17-JAN-2002;  
 Board of Regents, The University of Texas System (US); Williams,  
 Sanders R. (US); Rothermel, Beverly (US)

FEATURES  
 source Location/Qualifiers  
 1..2212  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 25..618  
 /note="named protein product"  
 /codon\_start=1  
 /protein\_id="CA023812.1"  
 /db\_xref="GI:18697052"  
 /translation="MEVDLDLPDPSATTACHLDPRVVDGLCRAKFSLFRTYDKDTT  
 FOYKSPKRVINFSNPLSADARLRLHKEFLGKMKLYFAOTLHGSSHLAPNPD  
 KQFLSPSPVGMKQVEDATPVINDLVLSLGGKEKELMAADTTTPSVVHY  
 CESDQKEEBEEMERMRPKKIIQTRPPTPIHLS"

BASE COUNT 588 a 455 c 529 g 640 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.74e-91 Length: 2212  
 Score: 1002.50 Matches: 191  
 Percent Similarity: 97.98% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 3  
 Query Match: 95.66% Indels: 1  
 DB: 6 Gaps: 1

US-09-782-953-3 (1-198) x AX365324 (1-2212)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHisLeuAspPro 20  
 DB 25 ATGAGAGAGGTGACCTGACGACCTGCCAGCGCACCATGCTGCACCTGGACCCG 84  
 OY 21 ArgValAlpheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 DB 85 CGCGTGTGTGTGACGCGCTGTGCGGGCCAAATTGAGTCTCTTAAAGACGATGAC 144  
 OY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheserAsn 60  
 DB 145 AAGACATCACCTTTCAGTATTTTAAAGCTTCAAAAGAGTCAAGATTAACCTTACAGCAAC 204

```
QY 61. ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisIysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTCTCCGCGAGATGCCAGGCTCCAGCTGCATGAAGACTGAGTTCTTGGGAAGGAA 264
QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 325 CCAGACAAGCAGTTCTTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTG 384
QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCAGTTCATGAATTTGACCATAGGAGTGAAGAGAAATGGAAAGATGAGGACCT 561
QY 141 GlyGluLysTyrGluLeuHisIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 562 AAGCCAAAATTTATCCAGACCGAGGCGGAGTACACGCCGATCCACCTCAGC 615

RESULT 6
HSU85266 2227 bp mRNA linear PRI 19-APR-2000
LOCUS Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
DEFINITION alternative exon 1, complete cds.
ACCESSION U85266
VERSION U85266.2 GI:7596913
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and Estivill,X.
TITLE A new human gene from the Down syndrome critical region encodes a
JOURNAL proline-rich protein highly expressed in fetal brain and heart
MEDLINE Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
PUBMED 96121593
REFERENCE 2 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M.A. and Estivill,X.
TITLE Genomic organization, alternative splicing, and expression patterns
JOURNAL of the DSCR1 (Down syndrome candidate region 1) gene
MEDLINE Genomics 44 (3), 358-361 (1997)
PUBMED 97468152
REFERENCE 3 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M. and Estivill,X.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca
MEDLINE Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
PUBMED 08907, Spain
REFERENCE 4 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M. and Estivill,X.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca
MEDLINE Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
PUBMED 08907, Spain
REMARK Sequence update by submitter
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612865.
FEATURES
Fuentes, J.J. Hum.Mol.Genet. 4, 1935-1944, 1995.
Location/Qualifiers
1..2227
/organism="Homo sapiens"

/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..2227
/gene="DSCR1"
25..618
/gene="DSCR1"
/note="down syndrome candidate region 1; one of four
alternatively spliced exon 1"
/codon_start=1
/db_xref="GI:7596914"
/translation="NEEVDLQPLSATIACHLDPRFVDGLCKRAKPFSLRTYDKDIT
FOYFKSRVRINFSNPFSAADARLQHKTEFLKEMKLYFAQTILHIGSHLAPNPD
KQFLISPPASPPVGHKQVEDATPVNYDLLVAISKLGPEKELYELHATDTTTSVVVHV
CESDQKEKEEBEEMRRPKLIIGRRPEYTPHLS"
BASE COUNT 603 a 455 c 529 g 640 t
ORIGIN

Alignment Scores:
Pred. No.: 8 81e-91 Length: 2227
Score: 1002.50 Matches: 191
Percent Similarity: 97.98% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 3
Query Match: 95.66% Indels: 1
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x HSU85266 (1-2227)
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 25 ATGGAGGAGTGGACCTGCGAGGACCTGCGAGCCGCCACCATCGCTGTACCTGGAGCCG 84
QY 21 ArgValPheValAspClyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CGCGTGTTCGTGGACGGCTGTGCGGGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 144
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGATAAATCTTCAGCAAC 204
QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGCGAGCAGATCCAGGCTCCAGTGCATAGACTGAGTTCTTGGGAAGGAA 264
QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 325 CCAGACAAGCAGTTCTTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTG 384
QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCAGTTCATGAATTTGACCATAGGAGTGAAGAGAAATGGAAAGATGAGGACCT 561
QY 141 GlyGluLysTyrGluLeuHisIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 445 GGGGAAAAGTATGAATTTGACCATAGGAGTGAAGAGAAATGGAAAGATGAGGACCT 561
QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTGCATCAGAG---AAGGAGGAGAGAGAGGAAATGGAAAGATGAGGACCT 561
QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 562 AAGCCAAAATTTATCCAGACCGAGGCGGAGTACACGCCGATCCACCTCAGC 615

RESULT 7
BC002864 BC002864 2289 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, Down syndrome critical region gene 1, clone MGC:10263
DEFINITION
```

ACCESSION IMAGE:3944959, mRNA, complete cds.  
 VERSION BC002864  
 KEYWORDS BC002864.1 GI:12804022  
 SOURCE MGC.

ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2289)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://imgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>

Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Shevchenko, Y., Wechterby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stancin, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNL at: <http://imgc.llnl.gov>  
 Series: IRAL Plate: 15 Row: a Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 7596913.

FEATURES  
 SOURCE Location/Qualifiers

1..2289  
 /organism="Homo sapiens"  
 /db\_xref="LocustID:1827"  
 /db\_xref="taxon:9606"  
 /clone="MGC:10263 IMAGE:3944959"  
 /cfeature\_type="Lung, small cell carcinoma"  
 /clone\_lib="NIH MGC 7"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOT87"  
 66..659  
 /codon\_start=1  
 /product="Down syndrome critical region gene 1"  
 /protein\_id="AAH02864.1"  
 /db\_xref="GI:12804023"  
 /translation="MEVDLQDLPSATTAACHDPRVFPVDFGCRAPFESLFRYDIT  
 FOYRFSFKRVRINFSNPASADARLQHLKTEFLGKEMLYEFAQTALHSSSHAPNP  
 KQFLISPAVPGKQVEDATPVNIQDLVAISLKGGEKEXKATLAADTTTSVVVAVH  
 CESDKEEEMERMRPKRIOTRPEYPIHLS"

CDS  
 BASE COUNT 628 a 465 c 555 g 641 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9.1e-91 Length: 2289  
 Score: 1002.50 Matched: 191  
 Percent Similarity: 97.98% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 3  
 Query Match: 95.66% Indels: 1  
 DB: 9 Gaps: 1

US-09-782-953-3 (1-198) X BC002864 (1-2289)

0y 1 MetGluValAlaPheLeuGlnAspLeuProSerAlaThrTlaAlaCywhbelauppro 20  
 |||

Db 66 ATGAGAGAGTGTGACCTGACGAGACCTGCCAGCGCCACCACTCGCTGTGACCTGGACCG 125  
 Oy 21 ArgValPheValaAPGlyLeuCyArgAlaLysPheGlySerLeuPheArgTh-TyrAsp 40  
 Db 126 CGCGTGTCTGTGACGCGCGCTGTGCGCGGCCAAATTATGATCCCTTTAGGACGTATGAC 185  
 Oy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValaArgLysPheSerAsn 60  
 Db 186 AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAACGAGTCAGAAATMACTTACGACAC 245  
 Oy 61 ProLeuSerAlaAlaAPAlaAPAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 Db 246 CCTTCTCCGACGACGAGATGCTCCAGCTCCAGCTGATAGATGATGTTCTTGGAAAGGAA 305  
 Oy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLysAlaProProAsn 100  
 Db 306 ATGAAGTATATTTTGTGCTCAGACCTTACACATAGGAAGCTCACCTGCGCCCAAT 365  
 Oy 101 ProAspLysGlnPheLeuLysSerProProLysSerProProValaGlyTyrLysGlnVal 120  
 Db 366 CCAGCAAGCAGTGTGTGATCTCCCTCCCTCCCTCCGCGGAGTGGAAACAAAGTG 425  
 Oy 121 GluAspAlaThrProValaLysAsnTyrAspLeuTyrAlaLysSerLysLeuGlyPro 140  
 Db 426 GAGATGCGACCCGACGATCATATGATCTTATATGTCATCTCAAGCTGGGCGCA 485  
 Oy 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValaValaHisVal 160  
 Db 486 GGGGAAAGTATGATGATTCACGACGCGCTGACACCACTCCAGCGTGTGTGCATCTA 545  
 Oy 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
 Db 546 TGTGAGATGATCAAGAC--AAGAGGAAAGAGGAAAGGAAAGGAAAGATAGAGGACCTT 602  
 Oy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
 Db 603 AAGCCAAATATTCACACAGAGCGGAGTACAGCGGATCCACTCAGC 656  
 RESULT 8  
 AK092184  
 LOCUS  
 DEFINITION Homo sapiens cDNA FLJ34855 fis, clone NT2NE2014104, highly similar  
 to DOWN SYNDROME CRITICAL REGION PROTEIN 1.  
 ACCESSION AK092184  
 VERSION AK092184.1 GI:21750714  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,  
 clone lib:NT2NE2 clone:NT2NE2014104.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Niimura, K., Magatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,  
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,  
 Kamihara, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M.,  
 Yamayama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
 Yanamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,  
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,  
 Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,  
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
 Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2407)  
 Isogai, T. and Yamamoto, J.  
 TITLE JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: [genomic@hri.co.jp](mailto:genomic@hri.co.jp), Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan

\* Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

# FEATURES

source  
1..2407  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NT2NE2014104"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2NE2"  
/note="Cloning vector: pME18SF13-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."

BASE COUNT 607 a 519 c 627 g 654 t

## ORIGIN

Alignment Scores:  
Pred. No.: 9,67e-91 Length: 2407  
Score: 1002.50 Matches: 191  
Percent Similarity: 97.98% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 3  
Query Match: 95.66% Indels: 1  
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AK092184 (1-2407)

Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
Db 218 ATGGAGAGGTGGACCTGCAGACCTGCCAGCGCCACCATCGCTGTCACCTGGACCCG 277  
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyxAsp 40  
Db 278 CGCGTGTTCGTGGAGCGCTGCCGCGCCAAATTTGAGTCCCTTTAGACCGATGAC 337  
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
Db 338 AAGACATCACCCTTCAGTATTTTAAGAGCTTCAACAGAGTCAGAAATAAACTTCAGCAAC 397  
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 398 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTCCTGGGAAGGAA 457  
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerHisLeuAlaProProAsn 100  
Db 458 ATGAAGTTATATTTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 517  
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120  
Db 518 CCAGACAGCAGTTTCGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAAACAGTG 577  
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 578 GAAGATCGACCCAGTCATAAATATGATCTCTTATATGCCATCTCAAGCTTGGGCCA 637  
Qy 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
Db 638 GGGAAAGATGATGAATTTGACGCGAGCAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 697  
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180  
Db 698 TGTGAGATGATCANGAG---AAGAGGAAGAGAGGAATGAAAGATGAGGACACCT 754  
Qy 181 LysProLysIleLeuGlnThrArgProGluTyrThrProLysHisLeuSer 198  
Db 755 AAGCAAAATATTCCAGACAGGAGCGGAGTACAGCGCGATCCACCTCAGC 808

## RESULT 9

AX365315  
LOCUS AX365315  
DEFINITION Sequence 5 from Patent W00204491.  
ACCESSION AX365315  
linear PAT 15-FEB-2002

## VERSION

AX365315.1 GI:18697045

## KEYWORDS

house mouse.

## SOURCE

Mus musculus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Williams, S.R. and Rothermel, B.

## AUTHORS

## TITLE

## JOURNAL

## Patent

## Board of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rothermel, Beverly (US)

## FEATURES

## source

1..597  
/organism="Mus musculus"

/db\_xref="taxon:10090"

## CDS

1..597  
/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD23809.1"

/db\_xref="GI:18697045"

/translation="MDRDFSYNFSLIACVANDVFSSETRAKFESLRTYDKDTT

FQYKSFKRVIRNFNPLSADARLRLHKLHTEFLGKMKLYPAQTLHGSSHLAPNPD

KQFLISPPASPVPVWKQVEDATPVNYDLVLAISKLGPKGKYLHAAATDPTFVVVHV

CESDQENEESEEMERMKPKKIQTTRRPTPIHLS"

BASE COUNT 170 a 156 c 142 g 129 t

## ORIGIN

## Alignment Scores:

Pred. No.: 2,95e-83 Length: 597  
Score: 920.50 Matches: 178  
Percent Similarity: 92.42% Conservative: 5  
Best Local Similarity: 89.90% Mismatches: 12  
Query Match: 87.83% Indels: 3  
DB: 6 Gaps: 1

US-09-782-953-3 (1-198) x AX365315 (1-597)

Qy 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspPro 20  
Db 1 ATGGATTTTAGGGACTTTAGCTACAAATTTTAGCTCCCTGATTTGCTGGCAACGAT 60  
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyxAsp 40  
Db 61 GATGCTTTCAGCGAAAGTGCAGACCGAGGCCAAATTTGAATTCCTCTCAGAACATATGAC 120  
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
Db 121 AAGGACACCACTTCCAGTATTTTAAGACTTCAACAGCTGTCGGGATAAATTCAGCAAC 180  
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 181 CCTTTATTCGACGCGCATGCCAGGCTGCCAGAGCCGAGTTCCTGGGAAGGAA 240  
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
Db 241 ATGAAGTTGTATTTTGTCTCAGACTTTACATAGGAAGTTACACTGGCTCCGCCCAAT 300  
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120  
Db 301 CCCGACAAACAGTCTCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGGAAACAGTA 360  
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 361 GAAGATGCCACCCCGTCATAAATTTACGATCTTTTATATGCCATCTCCAAGCTGGGCA 420  
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
Db 421 GGAGAGAGATGATGAATGCTGATGCGAGCGACGCCACCTCCAGTGTGTGTCCACGTG 480  
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180

Db	481	TGTGAGACTGACCAAGACATGAGAGAGAGACGACGAGAGATGAGAGAGATGAGAGATGAGAGACCC	540
QY	181	LYSPOLYSLIEIEGLThrArgArgProGluTyrThrProIleHisLeuSer	198
Db	541	AAGCCAAATCATCTCAGACACGGAGCCGGAGTACACACCGATCCACTTAC	594
RESULT 10			
AF237790			
DEFINITION	Mus musculus myocyte-enriched calcineurin interactor protein 1		
ACCESSION	AF237790		
KEYWORDS	splice variant 4 mRNA, complete cds.		
SOURCE	AF237790.1 GI:7542528		
ORGANISM	Mus musculus.		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 597)		
TITLE	Rothenberg, B.A., Vega, R.B., Yang, J., Wu, H., Baszel-Dubay, R.S. and Williams, R.S.		
JOURNAL	A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling		
REFERENCE	J. Biol. Chem. (2000) In press		
AUTHORS	2 (bases 1 to 597)		
TITLE	Rothenberg, B.A., Vega, R.B., Yang, J., Wu, H., Baszel-Dubay, R.S. and Williams, R.S.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
FEATURES			
SOURCE	1..597	Location/Qualifiers	
CDS	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="16"		
	1..597		
	/note="similar to the EST with GenBank Accession Number AA200984; similar to Homo sapiens DSCR1"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin interactor protein		
	1 splice variant 4"		
	/protein_id="AA063486.1"		
	/db_xref="GI:7542529"		
	/translation="MDPRDFSYNFSLLACVANDVFSSETRAKFESLFRTRYDKDTT		
	FOYFSPKRVIRINSPSAADARLRLHTEPELGEKMKLYFQTLHIGSSHLAPNPPI		
	KQFLSPASPVGKQVNDAPVNIYDLIVLSLQPEKEXELHAATDPFSPVYVH		
	KQFLSPASPVGKQVNDAPVNIYDLIVLSLQPEKEXELHAATDPFSPVYVH		
	CESDQSENEEEEMERMKRPKRIITRTPEPTPIHLS"		
BASE COUNT	170 a 156 c 142 g 129 t		
ORIGIN			
Alignment Scores:			
Prid. No.:	2,95e-83	Length:	597
Score:	920.50	Matches:	178
Percent Similarity:	92.42%	Conservative:	5
Best Local Similarity:	89.90%	Mismatches:	12
Query Match:	87.83%	Indels:	3
DB:	10	Gaps:	1
US-09-782-953-3 (1-198) x AF237790 (1-597)			
QY	4	VALASPLEGGINAPLEUPRO-----SERALATHRIIEALCYSHISLEAEP	20
Db	1	ATGATTTTGGGACTTTAGCTACATTTTAAAGTCCCTCGATTCGTTGTGTGGCAAGAT	60
QY	21	ARGVALPHEVALAAGPGYLLECYARGALALYSPHEGIGSERLEUPHEARGTHTYASP	40
Db	61	GATGCTTTCAGCGAAGAGAGACACGAGCCCAATTTAAATCCCTCTTCAGACATATGAC	120
QY	41	LYSPHTHTHTHPAGINTYRPHYLSESERPHELYSARGVALARGILEANPSESERSN	60
Db	121	AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAACAGTGGCCGATAACTTCAGCAAC	180

[illegible]

Percent Similarity: 94.68%  
 Best Local Similarity: 93.09%  
 Query Match: 86.93%  
 DB: 10

Conservative: 3  
 Mismatches: 10  
 Indels: 0  
 Gaps: 0

US-09-782-953-3 (1-198) x AF263239 (1-2141)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
 DB 61 AGCTCCCTGATGTTGTTGTCGCAACGATGATGTTCTTCAGCGAAGTGACGACGGCC 120  
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
 DB 121 AAATTGAATCCCTTCAGACACATATGACAAAGACACACCTTCCTCCAGTATTTAAGAGC 180  
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 DB 181 TTCAAAGTGTCGCGATAAATTCAGCAACCCCTTATCTGACGCGATGCGACGCTCGG 240  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 DB 241 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 300  
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110  
 DB 301 ATAGGAAGTTCACACTGCTGCTCCGCCCAATCCCGACAAACAGTTCCTCATCTCCCTCCG 360  
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
 DB 361 GCCTCTCTCCCTGCTGGTGGAAACAAGTAGAGAGTCCACCCCGTCATAAATTACGAT 420  
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysTyrGluLeuHisAlaAlaThr 150  
 DB 421 CTTTATATGCACTCTCAAGCTGGGGCCAGGAGAGAATGAACTGCATGCGACGACA 480  
 QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
 DB 481 GACACCACTCCAGTGTGTGGTCCAGCTGTGAGAGTACCAAGAGATGAGGAGAA 540  
 QY 171 GluGluMetGluArgMetLysArgProLysProLysIleLeuGlnThrArgArgPro 190  
 DB 541 GAGGAAGAGATGGAGAGAATGAAGAGACCAAGCCCAAAATCATCCAGACAGGACCG 600  
 QY 191 GluTyrThrProIleHisLeuSer 198  
 DB 601 GAGTACACCACTCCACCTCAGC 624

BC013551 2198 bp mRNA linear ROD 07-AUG-2002  
 LOCUS Mus musculus, Down syndrome critical region homolog 1 (human),  
 clone MGC:19348 IMAGE:4236038, mRNA, complete cds.

ACCESSION BC013551  
 VERSION BC013551.1 GI:15488840  
 KEYWORDS MGC.

SOURCE Mus musculus  
 ORGANISM house mouse.

REFERENCE 1 (bases 1 to 2198)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-SEP-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 25 Row: m Column: 6.

## FEATURES

Location/Qualifiers  
 1..2198  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="FVB/N"  
 /clone="MGC:19348 IMAGE:4236038"  
 /tissue\_type="kidney, normal. 5 month old male mouse."  
 /clone\_lib="NCI CGAP\_Kid14"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 77..673  
 /codon\_start=1  
 /product="Down syndrome critical region homolog 1 (human)"  
 /protein\_id="AAH13551.1"  
 /db\_xref="GI:15488841"  
 /db\_xref="LocusID:54720"  
 /translation="MHRDFSYNFSLLIACVANDVDFSETRAKFESLFRITYDKDTT  
 FOYPKSKVRINFSNPLSADARLRLHKTFELKEMKLYFAOTLHIGSSHLAPNPDP  
 KQFLSPASPPGVKQVEDATPVINYDLLTAISKLGEGKEYELHAATDTTPSVVVHV  
 CESDNEBEEBEMERMKRPKPIITRRPEYTHLS"  
 BASE COUNT 533 a 539 c 569 g 557 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,268-81 Length: 2198  
 Score: 911.00 Matches: 175  
 Percent Similarity: 94.68% Conservative: 3  
 Best Local Similarity: 93.09% Mismatches: 10  
 Query Match: 86.93% Indels: 0  
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x BC013551 (1-2198)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
 DB 107 AGCTCCCTGATGTTGTTGTCGCAACGATGATGTTCTTCAGCGAAGTGACGACGGCC 166  
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
 DB 167 AAATTGAATCCCTTCAGACACATATGACAAAGACACACCTTCAGTATTTAAGAGC 226  
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 DB 227 TTCAAAGTGTCGCGATAAATTCAGCAACCCCTTATCTGACGCGATGCGACGGCC 286  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 DB 287 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 346  
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110  
 DB 347 ATAGGAAGTTCACACTGCTGCTCCGCCCAATCCCGACAAACAGTTCCTCATCTCCCTCCG 406  
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
 DB 407 GCCTCTCTCCCTGCTGGTGGAAACAAGTAGAGATGCGACCCCGCTCATAAATTACGAT 466  
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysTyrGluLeuHisAlaAlaThr 150  
 DB 467 CTTTATATGCACTCTCAAGCTGGGGCCAGGAGAGAATGAACTGCATGCGACGACA 526  
 QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
 DB 527 GACACCACTCCAGTGTGTGGTCCAGCTGTGTGAGAGTGAACCAAGAGATGAGGAGAA 586



OY	171	GIGUGUJMEGLUARGMELVYARCPOLYSPPOLYSLEILEGILTHARGARFP	190
Db	587	GAGGAAGATGAGAGANTGAAAGAACCCCAAGCCCAAATCATCTCACAACGGAGACG	646
OY	191	GIUTYTRHPRIOLEHIEIeUser	198
Db	647	GAGTACACACCCTCACCTCAGC	670
RESULT 13			
LOCUS	AF263240	2224 bp	mRNA linear ROD 30-MAY-2000
DEFINITION	Mus musculus calcineurin inhibitor mRNA, complete cds,		
VERSION	AF263240		
KEYWORDS	AF263240.1 GI:8102013		
SOURCE			
ORGANISM	Mus musculus.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2224)		
AUTHORS	Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.		
TITLE	Down syndrome candidate region 1 (Dscr1), one of three		
JOURNAL	alternatively spliced exon 1 transcripts		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2224)		
TITLE	Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.		
JOURNAL	Direct Submission		
	Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,		
	Avila, Castelldefels Km. 2,7, L'Hospital del de Liobregat, Barcelona		
	08907, Spain		
FEATURES			
source	Location/Qualifiers		
CDS	1..2224		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="16"		
	354..710		
	/note="Down syndrome candidate region 1; Dscr1;		
	alternatively spliced exon 1"		
	/product="calcineurin inhibitor"		
	/protein_id="AAF72702.1"		
	/db_xref="GI:8102014"		
	/translation="MKLYPAQTIHTIGSSHLAPNDPKFLISPPASPYGVKQVEDAI		
	PYNVDLVAISKLGGEKXELHAATDTTPSVVHVCSDDQNEEBEEMERMKPKIK		
	KIIOTRRPEYPIHLS"		
BASE COUNT	519 a 554 c 586 g 562 t	3 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	1,61e-79	Length:	2224
Score:	890.00	Matches:	169
Percent Similarity:	99.41%	Conservative:	0
Best Local Similarity:	99.41%	Mismatches:	1
Query Match:	84.92%	Indels:	0
DB:	10	Gaps:	0
US-09-782-953-3 (1-198) x AF263240 (1-2224)			
OY	29	ATGTAAlayPhegiUsSerLeuPheArgrThTyTraspiyAsePhrThRphesgIntyrPhe	48
Db	198	AGGCCCAAAATTGTGATCCCTTCAGAACATATGACAAGAAGACACACTTCAGATATT	257
OY	49	LysSerPheLySAyGVALArgIleAsnPhSeSerAnPrOleuSerAlaIAeSpALArg	68
Db	258	AAGAGCTTCAAACCTGTCGGATAAATTCAAGCAACCCCTTATCTGAGCGAGGCAGG	317
OY	69	LeuArgrLeuHiSLyBThrGUPhLeuGUlyLySGluWeLySeuTYrPheAlagInThr	88
Db	318	CTGGGCTGCACAGACGAGCTTCTCGGGAGAGAAATGAAGITGTATTTTGCTCAACT	377
OY	89	LeuHiSLieglSerSerHiLeuAlarProForoanPrOleapLySGInPheLeuIleSer	108

D6	TTACACTAGGAAGTTCACACCTGGGCTCCGCCCAATCCCGACAACAGATTCTCATCTCC	437
OY	109 ProProAlaSerProProValGIYTPVYGInvalGIUAspAlaThrProValIleLeu	128
D6	438 CCTCGGCCTCTCCTCCCGTGCTGGAAACAAGTGAAATATCCACCCCCTCATTAAT	497
OY	129 TyrAspLeuLeuTYFAIaIIeSerLYLeuGLYProGLYQIUubTYSTYrGIueLHIAIa	148
D6	498 TTAGACTTTTATATAGCCATCTCCAACTGGGGCCAGAGAAGAAAGTATGAACCTGCATGCA	557
OY	149 AlaThrAspProThrProSerValValIaHisValCYseGIuSerApgIngluaenglu	168
D6	558 GCGACAGACACCCTCCCAAGTGTGTGTGCACGTGTGTAGAGTACCAAGAGAATGAG	617
OY	169 GIUGUGUGUGUGUMerGIUAyMeLIyARgProLySryolylellellegINTTrArg	188
D6	618 GAGGAGAGGAGAAAGATGAGAGATGAAGAACCCCAAGCCAAATATCATCCAGACACGG	677
OY	189 ArgProGIUTYrThrProIleHIAleuSer	198
D6	678 AGACCGAGTACACACCCATCCACTCACG	707
RESULT 14		
LOCUS	AB075973	626 bp mRNA linear ROD 27-JUL-2002
DEFINITION	Rattus norvegicus mcip 1 mRNA for myocyte-enriched calcineurin-interacting protein 1, complete cds.	
ACCESSION	AB075973	
VERSION	AB075973.1	GI:21998843
KEYWORDS		
SOURCE	Rattus norvegicus CDNA to mRNA.	
ORGANISM	Rattus norvegicus Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1	
REFERENCE	Odashima,M., Nagata,K., Obata,K., Somura,F., Izawa,H., Miyazaki,T., Murata,Y., Seo,H., Yamada,Y. and Yokota,M. rat myocyte-enriched calcineurin interactin protein 1, splice variant 4 mRNA, complete cds Unpublished 2 (bases 1 to 626)	
JOURNAL	Obata,K., Nagata,K., Odashima,M., Somura,F., Yamada,Y. and Yokota,M. Direct Submission Submitted (11-DEC-2001) Koji Obata, Nagoya University, Department of Clinical Pathophysiology; 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan (E-mail:yobata@surui.med.nagoya-u.ac.jp, Tel:81-52-744-2577, Fax:81-52-744-2977)	
AUTHORS		
TITLE	Location/Qualifiers	
FEATURES		
source	1..626 /organism="Rattus norvegicus" /db_xref="taxon:10116"	
gene	1..626 /gene="mcip 1"	
CDS	19..618 /gene="mcip 1" /codon_start=1 /product="myocyte-enriched calcineurin-interacting protein 1" /protein_id="BAC06443.1" /db_xref="GI:21998843" /translation="MHPRDDPVNFSSLIACVANGDVFSESTRAPKPSLPFYDKDTIT FOYSSEFRVRINFNSNPLSAADALRLHKTFPELOKEMLKYAQOLIHGSSLAPPND KQFLSPASPVGMKVDEATPIVINDLPAISKLGEKRIELHAATDTTPSYVVHV CESDOEEEEEEMERMKRPKITIGTRREYPIHLIS"	
BASE COUNT	182 a 163 c 154 g 127 t	
ORIGIN		
Alignment Scores:	6,286-80 Length: 626	
Pred. No.:	887,50 Matches: 173	



Percent Similarity: 93.12% Conservative: 3  
 Best Local Similarity: 91.53% Mismatches: 12  
 Query Match: 84.69% Indels: 1  
 DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x AB075973 (1-626)

QY 11 SerAlaThrIleAlaCysHisLeuAspProHrgValPheValAspGlyLeuCysArgAla 30  
 DB 49 AGCTCCCTGATGCTGTGGCAACCGGTGATGCTTTCAGCGAAGTGAACCGAGGCC 108  
 QY 31 LysPheGluSerLeuPheAsgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50  
 DB 109 AAATTTGAATCCCTTTCAGGAGCTATGACAGGACATCACCTTCCAGTATTTTAAAGC 168  
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 DB 169 TTCAAGCGTGTCCGAATAAACTTCAGCAACCCCTGCTGCGGCGGAGCGGCTCGG 228  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 DB 229 CTGCATAGACAGAGTTCTCTGGGGAAGGAGATGAATGTAATCTTTGCCAGACTTTACAC 288  
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110  
 DB 289 ATAGGAAGTTCAACCTGGCTCGCCCAATCCAGCAACAGATTCTCATCTCTCCCCCT 348  
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
 DB 349 GCCTCTCCACCTGTGTGGTGGAAACAGTAGAGAGCTACCCAGTCATAAACTAGCAT 408  
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
 DB 409 CTTTATATCTATCTCAAGCTGGGACCGAGGAGAGTACGAGCTACACGAGCGACA 468  
 QY 151 AspProThrProSerValValHisValCysGluSerAspGln---GluAsnGluGlu 169  
 DB 469 GACACACGCCCGCGTGTGTCCAGTGTGTGAGAGTACCAAGAGGAGGAGGAGAA 528  
 QY 170 GluGluGluMetGluArgMetLysArgProLysProlysIleIleGlnThrArgArg 189  
 DB 529 GAGGAGGAGGAGTGGAGAGAATGAAGAGACCCCAAGCAAAATCATCCAGACTCGAGG 588  
 QY 190 ProGluTyrThrProIleHisLeuSer 198  
 DB 589 CCAGAGTACACCCATCCACCTCAGC 615

RESULT 15  
 CGU60263 2216 bp mRNA linear ROD 03-SEP-1997  
 LOCUS  
 DEFINITION Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.  
 ACCESSION U60263  
 VERSION U60263.1 GI:2351390  
 KEYWORDS  
 SOURCE Cricetulus griseus.  
 ORGANISM Cricetulus griseus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 Cricetulus.  
 1 (bases 1 to 2216)  
 Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and  
 Davies,K.J.  
 TITLE Hamster adapt78 mRNA is a Down syndrome critical region homologue  
 that is inducible by oxidative stress  
 JOURNAL Arch. Biochem. Biophys. 342 (1), 6-12 (1997)  
 MEDLINE 97329095  
 PUBMED 9185608  
 REFERENCE 2 (bases 1 to 2216)  
 Crawford,D.R., Leahy,K.L. and Davies,K.J.A.  
 AUTHORS Direct Submission  
 TITLE Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical  
 JOURNAL College, 47 New Scotland Avenue, Albany, NY 12208, USA  
 FEATURES Location/Qualifiers

## source

1..2216  
 /organism="Cricetulus griseus"  
 /db\_xref="taxon:10029"  
 /cell\_type="HA-1"

## gene

1..2216

/gene="adapt78"  
 70..663  
 /gene="adapt78"  
 /note="similar to human DSCR1; an oxidant- and  
 calcium-inducible mRNA"

## CDS

/codon\_start=1  
 /product="Adapt78"  
 /protein\_id="AAB68517.1"  
 /db\_xref="GI:2351391"  
 /translation="MHFRDENVFSLIACVANGDVSESETRAFESLFRTYDKDIT  
 FOYKSFYKVRINFENPISAADARLQLHKTEFLGKEMKLYPAQTLLHGSSHLAPPND  
 KQFLSPSPASPVGWKQVEDATPVINYLLEYAISKLGPGEKLYELHAATDTTPSVVVHV  
 KESLDENEEEMERMKRPKPKIITRPEYTPHLS"

BASE COUNT 544 a 515 c 564 g 593 t

## ORIGIN

## Alignment Scores:

Prod. No.: 5,71e-79 Length: 2216  
 Score: 884.50 Matches: 172  
 Percent Similarity: 93.62% Conservative: 4  
 Best Local Similarity: 91.49% Mismatches: 11  
 Query Match: 84.40% Indels: 1  
 DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x CGU60263 (1-2216)

QY 11 SerAlaThrIleAlaCysHisLeuAspProA:GValPheValAspGlyLeuCysArgAla 30  
 DB 100 AGCTCCCTGATGCTGTGGCAACCGGTGATGCTTTCAGCGAAGTGAACCGAGGCC 159  
 QY 31 LysPheGluSerLeuPheAsgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50  
 DB 160 AAATTTGAATCCCTTTCAGGAGCTATGACAGGACATCACCTTCCAGTATTTTAAAGC 219  
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 DB 220 TTCAAGCGTGTCCGAATAAACTTCAGCAACCCCTTATCCGAGCTACCCAGGCTCGAG 279  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 DB 280 CTGCATAGACAGGAGTTCTCTGGGGAAGGAAATGAAGCTGTACTTTGCTCAGAGTTACAC 339  
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110  
 DB 340 ATAGGAAGTTCAACCTGGCTCGCCCAATCCAGCAACAGATTCTCATCTCTCCCTCC 399  
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
 DB 400 GCCTCTCCACCAAGTGTGTGGAGCAAGTAGAAGAGCTTACCCAGTATAAATACGAT 459  
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
 DB 460 CTTTATATGCTATCTCAAGCTGGGCGGAGGTGAGAGTATGAGCTGATCGATCGACGACA 519  
 QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170  
 DB 520 GACACCACTCCAGTGTGTGAGTCCACCGTGTGTGAAGCGACCAAGAGAAT---GAAGAG 576  
 QY 171 GluGluGluMetGluArgMetLysArgProLysProlysIleIleGlnThrArgArgPro 190  
 DB 577 GAGGAGGAGATGGAGAGATGAAGAGACCCCAAGCAAAATTTATCCAGACGAGGCCA 636  
 QY 191 GluTyrThrProIleHisLeuSer 198  
 DB 637 GAGTACACGCTTATCCACCTCAGC 660

Search completed: December 14, 2002, 21:31:58

Sun Dec 15 08:38:11 2002

us-09-782-953-3.rge

Job time : 1660.79 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3457 Seconds  
(without alignments)  
1117.976 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEESVDLQDLPSAIIACHLDP.....RPRKLIQTRRPTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	99.2	198	1	CCP1_MOUSE
2	1002.5	95.7	197	1	CCP1_HUMAN
3	884.5	84.4	197	1	CCP1_CRIGR
4	613.5	58.5	197	1	CCP2_MOUSE
5	597	57.0	239	1	CCP3_MOUSE
6	596.5	56.9	197	1	CCP2_HUMAN
7	596	56.9	241	1	CCP3_HUMAN
8	379.5	36.2	292	1	NLA_DROME
9	351	33.5	207	1	CCPL_CABEL
10	95.5	9.1	163	1	YA9F_SCHPO
11	88	8.4	630	1	YCF2_OENVI
12	88	8.4	721	1	YCF2_OENPI
13	86.5	8.3	822	1	NAH1_CRIGR
14	86	8.2	338	1	GALE_MYCPN
15	83	7.9	561	1	SNTC_HUMAN
16	82.5	7.9	827	1	MOF_DROME
17	81.5	7.8	1790	1	SEPA_EWEMI
18	81	7.7	150	1	ALL5_HEVBR
19	80.5	7.7	280	1	UPPS_METJA
20	80.5	7.7	360	1	VTPX_TTV1
21	79.5	7.6	904	1	DLG1_HUMAN
22	79	7.5	1142	1	JAK1_HUMAN
23	78.5	7.5	211	1	RCN1_YEAST
24	78	7.4	467	1	INVO_MOUSE
25	77.5	7.4	627	1	DNAK_GUITH
26	77	7.3	1033	1	S190_YEAST
27	77	7.3	1912	1	CHDA_HUMAN
28	76.5	7.3	558	1	ORC2_XENLA
29	76	7.3	445	1	HH3R_MOUSE
30	76	7.3	445	1	HH3R_RAT
31	76	7.3	450	1	INVO_LEMCA
32	76	7.3	1480	1	PANI_YEAST
33	75.5	7.2	594	1	YKA5_YEAST

## RESULT 1

### CCP1\_MOUSE

ID CCP1\_MOUSE STANDARD; PRT; 198 AA.  
AC Q9JHG6; Q9JKK3; Q9JKK2; Q9JK51; Q9JK50;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)  
DE (Myocyte-enriched calcineurin interacting protein 1) (MCIP1).  
GN DSCR1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RC STRAIN=BALB/c; TISSUE=Brain;  
RX MEDLINE=20534792; PubMed=11080588;  
RA Strippoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;  
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family;  
RT conserved synteny with the human orthologous genes.";  
RL Gene 257:223-232(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).  
RX MEDLINE=20187590; PubMed=10722714;  
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,  
RA Williams R.S.;  
RT "A protein encoded within the Down syndrome critical region is  
RT enriched in striated muscles and inhibits calcineurin signaling.";  
RL J. Biol. Chem. 275:8719-8725(2000).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.  
RC TISSUE=Fetal brain;  
RX MEDLINE=21152920; PubMed=11231093;  
RA Casas C., Martinez S., Pritchard M.A., Fuentes J.J., Nadal M.,  
RA Guimera J., Arbones M., Florez J., Soriano E., Estivill X.,  
RA Alcantara S.;  
RT "Dscr1, a novel endogenous inhibitor of calcineurin signaling, is  
RT expressed in the primitive ventricle of the heart and during  
RT neurogenesis.";  
RL Mech. Dev. 101:289-292(2001).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORMS B AND C).  
RA Fuentes J.J., Pritchard M.A., Pucharcos C., Estivill X.;  
RT "Down syndrome candidate region 1 (Dscr1), one of three alternatively  
RT spliced exon 1 transcripts.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM C).  
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;  
RX MEDLINE=21085560; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

Q11186 caenorhabdi  
P13834 oryctolagus  
Q09293 caenorhabdi  
Q92180 bos taurus  
Q92974 homo sapien  
Q42877 lycopersico  
O88573 mus musculus  
P40318 saccharomyc  
Q94966 homo sapien  
P22105 homo sapien  
P40366 saccharomyc  
P39770 drosophila

## ALIGNMENTS

[illegible]

```

DN 15-OCT-2001 (Ref. 40, last sequence update)
DT 15-OCT-2001 (Ref. 40, last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-
DS enriched calcineurin interacting protein 1) (MC1P1) (Adapt78).
GN DSCR1 OR DSCR OR ADAPT78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxId:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RX MEDLINE=96121593; PubMed=8595418;
RA Fuentes J.-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
RA Escivill X.;
RT "A new human gene from the Down syndrome critical region encodes a
RT proline-rich protein highly expressed in fetal brain and heart.";
RL Hum. Mol. Genet. 4:1935-1944(1995).
RN [2]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=97468152; PubMed=9325060;
RA Fuentes J.-J., Pritchard M.A., Escivill X.;
RT "Genomic organization, alternative splicing, and expression patterns
RT of the DSCR1 (Down syndrome candidate region 1) gene.";
RL Genomics 44:358-361(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=lung;
RA Strausberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
RC TISSUE=Breast;
RA Crawford D.R., Leahy K.P., Davies K.J.A.;
RT "Adapt78, a calcium and oxidant-inducible RNA.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20320698; PubMed=10861295;
RA Fuentes J.-J., Genesca U., Kingsbury T.J., Cunningham K.W.,
RA Perez-Riba M., Bativill X., de la Luna S.;
RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RT calcineurin-mediated signaling pathways.";
RL Hum. Mol. Genet. 9:1681-1690(2000).
RN [6]
RP FUNCTION: Inhibits calcineurin-dependent transcriptional responses
RN by binding to the catalytic domain of calcineurin A. Could play a
RN role during central nervous system development.
RN [7]
RP ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
RN PRODUCED BY ALTERNATIVE SPLICING.
RN [8]
RP TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
RN muscle. Also expressed in all other tissues.
RN [9]
RP SIMILARITY: By calcium.
RN [10]
RP SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
RN [11]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.ebi.ac.uk/announcement/
RN or send an email to license@ebi-sib.ch).
RN [12]
RX EMBL, U28833; AAB81557.1; -.
DR EMBL, U85265; AAB84370.1; -.
DR EMBL, U85266; AAB84371.2; -.
DR EMBL, U85267; AAB84372.1; -.
DR EMBL, BC002864; AAH02864.1; -.
DR EMBL, U53821; AAF21218.1; -.
DR Genew; HGNC:3040; DSCR1.
DR MIM: 602917; -.
KW Alternative splicing
FT DOMAIN 166 174 POLY-GLU.
FT VARSPPLIC 1 28 MEVDLQDPATNCHLDPRYFUNGLC -> MPRRNPVNS
FT FSSVLLQVANSDFSESET (IN ISOFORM 2).

```

```
FT VARSPLIC 1 29 MEEVDQLDLPSTACHLDPRVFDGLCR -> MVY (IN
FT VARSPLIC 1 80 MISSING (IN ISOFORM 4).
FT CONFLICT 159 159 H -> Q (IN REF. 4).
SQ SEQUENCE 197 AA; 22637 MW; 1BD426BB88167E6C CRC64;

Query Match 95.7%; Score 1002.5; DB 1; Length 197;
Best Local Similarity 96.5%; Pred. No. 2e-80;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDQLDLPSTACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSN 60
DB 1 MEEVDQLDLPSTACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSN 60

QY 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPSPVGMKQV 120
DB 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPSPVGMKQV 120

QY 121 EDATPVINDLYLVAISKLGPEKVELHAATDTPSVVHVHVCESQENEEEMERMKRP 180
DB 121 EDATPVINDLYLVAISKLGPEKVELHAATDTPSVVHVHVCESQENEEEMERMKRP 180

QY 181 KPILQTRRPETPIHLS 198
DB 180 KPILQTRRPETPIHLS 197

RESULT 3
CCPI CRIGR STANDARD; PRT; 197 AA.
AC O35847;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Oxidative-induced protein Adapt78).
GN DSCR1 OR ADAPT78.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97329095; PubMed=9185608;
RX Crawford D.R., Leahy K.P., Abranava N., Lan L., Wang Y., Davies K.J.;
RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue
RT that is inducible by oxidative stress.";
RL Arch. Biochem. Biophys. 342:6-12(1997).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -!- INDUCTION: OXIDANT-INDUCIBLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U60263; AAB68517.1; -.
SQ SEQUENCE 197 AA; 22674 MW; 0EFC68FA55826439 CRC64;

Query Match 84.4%; Score 884.5; DB 1; Length 197;
Best Local Similarity 91.5%; Pred. No. 3.8e-70;
Matches 172; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 11 SATIACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSNPLSAADARLR 70
DB 11 SSLIACVANGDVSESTRAKFESLRTYDKDTTFQYFKSKRVIRNFSNPLSAADARLRQ 70
```

```
QY 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPSPVGMKQVEDATPVIND 130
DB 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPSPVGMKQVEDATPVIND 130

QY 131 LLVAISKLGPEKVELHAATDTPSVVHVHVCESQENEEEMERMKRPKPIQTRRP 190
DB 131 LLVAISKLGPEKVELHAATDTPSVVHVHVCESQEN-EEEEEMERMKRPKPIQTRRP 189

QY 191 EYTPIHLS 198
DB 190 EYTPIHLS 197

RESULT 4
CCP2 MOUSE STANDARD; PRT; 197 AA.
AC Q9JHG2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)
DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).
GN DSCR1l1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20187590; PubMed=10722714;
RX Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RT enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=20534792; PubMed=11080588;
RA Striappoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RT conserved synteny with the human orthologous genes.";
RL Gene 257:223-232(2000).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC and brain. Lower expression in all other tissues.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF237791; AAF63487.1; -.
DR EMBL; AF237887; AAF62538.1; -.
DR MGD; MGI:1858219; Dscr1l1.
SQ SEQUENCE 197 AA; 22025 MW; CEL306B7B03E70F1 CRC64;

Query Match 58.5%; Score 613.5; DB 1; Length 197;
Best Local Similarity 63.3%; Pred. No. 1.5e-46;
Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

QY 11 SATIACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSNPLSAADARLR 70
DB 11 STLVCACVDVFTVQEVKEKEGLEFRTYDECVTLQFLFKSFRVRINFSPKSAARARIE 70

QY 71 LHKTEFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPSPVGMKQVEDATPV 126
```

```

Db 71 LHEQFRGKLUKLYPAQVOTPEITDGLHLAPPOPAQFLISPPSPSPVGNKPPISDAPV 130
QY 127 INYDLVIAISKLGPEKEKELHAATDPTPSVYVHVCESEQDNEEBEEMERMRKPPKIIQ 186
Db 131 LNYDLVIAVAKLGPKEKELHAGTSTPSVYVHVCESEDEEDPK-----TSPKPIIQ 185
QY 187 TRREPTP 194
Db 186 TRREPLP 193

RESULT 5
CCP3_MOUSE STANDARD; PRT; 239 AA.
ID CCP3_MOUSE 09CX87;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
DE DSCR1L2.
CN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/c; TISSUE=Brain;
RA Strittopoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like gene family."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 40-239 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Baralov S., Casavant T.,
RA Fletschman W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bona M.P.,
RA Brownstein M.O., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombarer P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontuski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF237888; AAF62539.1; -
CC DR EMBL: AK019377; BAB31687.1; -
CC MGD: MG1:1658220; DSCR1L2.
CC DR MGD: MG1:1658220; DSCR1L2.
CC SEQ SEQUENCE 239 AA; 27153 MW; 1B2687B47B4D272 CRC64;

```

```

Query Match 57.0%; Score 597; DB 1; Length 239;
Best Local Similarity 60.7%; Pred. No. 5, 1e-45;
Matches 122; Conservative 20; Mismatches 49; Indels 10; Gaps 2;

QY 2 EENVLDLPSTATICHDPVFDGLCRAPFESLPRTYDQTFQYKSPKRVIRINSP 61
Db 38 EMDLSDPLTSLFACSVHEAVFEVQEKERFALFTLYDOVTFOLFKSRVRIRINSP 97
QY 62 LSAADARLRHKEFLKEMKLYFAQTLMIG-----SSHLAPPNDKQFLISPPASPVGW 117
Db 98 EAAARARIELHSEFHRKLUKLYPAQVQSGEARDKSYLLPPOPTKQFLISPPASPVGW 157
QY 118 KQVEDATPVINYDLVIAISKLGPEKEKELHAATDPTPSVYVHVCESEQDNEEBEEMER 177
Db 158 KQVEDATPVINYDLVIAISKLGPEKEKELHAGTSTPSVYVHVCESEDEEDPK-----T 211
QY 178 KRPKPIQTRREPTPILHS 198
Db 212 KRPKPIQTRREPTPILHS 198

RESULT 6
CCP2_HUMAN STANDARD; PRT; 197 AA.
ID CCP2_HUMAN 014206;
AC 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcipressin 2 (Thyroid hormone-responsive protein ZAKI-4) (Down
DE syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE interacting protein 2) (MCIP2).
DE DSCR1L1 OR ZAKI4.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=9627929; PubMed=8662924;
RA Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA Yamamura H., Seo H.;
RT "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT in human skin fibroblasts."
RL J. Biol. Chem. 271:14567-14571(1996).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -1- TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC and skeletal muscle but not in placenta, lung, kidney and
CC pancreas.
CC -1- INDUCTION: By thyroid hormone.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D83407; BA11911.1; ALT_INIT.
CC DR EMBL: HGNC:3041; DSCR1L1.
CC DR MIM: 604876; -
CC SEQ SEQUENCE 197 AA; 21955 MW; AF39735F6661CBED CRC64;

Query Match 56.9%; Score 596.5; DB 1; Length 197;
Best Local Similarity 62.2%; Pred. No. 4, 4e-45;
Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

QY 11 SATIACHLDPVFDGLCRAPFESLPRTYDQTFQYKSPKRVIRINSPISADALR 70

```

```

Db 11 STLVACVVDVEFTQVKEKPGGLFRTYDCVTFQLFKSPRRVRINFSNPKSARARIE 70
Qy 71 LKTEFLGKEMKLYFAQTLHGSS-----HLAPPNDKQFLISPPASPPVGVKQVEDATPV 126
Db 71 LHETOFGRGKLLKLYFAQVQTPETDGLHLAPPOPAKQFLISPPSSPPVQPPINDATPV 130
Qy 127 INVDLLYALSKLPGCEKVELHAADTPSPVVHVHVESDOENEEEMERMKRKPPIIQ 186
Db 131 LNYDILYAVAKLPGCEKVELHAGTSTSPVVHVHVCDSIDEEEDPK-----TSPKPIIQ 185
Qy 187 TRRPEYTP 194
Db 186 TRRGLPP 193

```

## RESULT 7

```

ID_CCP3 HUMAN STANDARD; PRT; 241 AA.
AC O9UKA8; O9UKA7; O9UC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN DSCR1L2.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Peripheral blood, and Placenta;
RX MEDLINE=2021370; PubMed=10756093;
RT Strippoli P., Lenzi L., Petrini M., Carinci P., Zannotti M.;
RT "A new gene family including DSCR1 (Down syndrome candidate region 1)
RT and ZAK1-4: characterization from yeast to human and identification of
RT DSCR1-like 2, a novel human member (DSCR1L2).";
RL Genomics 64:252-263(2000).
RN [2]
RP SEQUENCE OF 66-241 FROM N.A.
RA Bagguley C.;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (SHOWN HERE) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC kidney, liver and peripheral blood leukocytes. Lower expression in
CC all other tissues.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL; AF176116; AAF01684.1; -
CC DR EMBL; AF176117; AAF01685.1; -
CC DR EMBL; AL034582; CAB72298.1; -
CC DR Genew; HGNC:3042; DSCR1L2.
CC DR MIM; 605860; -
CC KW Alternative splicing.
CC FT VARSPLOC 124 133 MISSING (IN ISOFORM 2).
CC SQ SEQUENCE 241 AA; 27492 MW; 91AB619F5E443FBD CRC64;

```

```

Query Match 56.9%; Score 596; DB 1; Length 241;
Best Local Similarity 61.9%; Pred. No. 6.2e-45;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

```

```

Qy 2 EVDLDLPATLACHLDPRVPVDGLCRAKFESLFRDYKDTTFQYFKSPKRVINFSNP 61
Db 38 EMDLSDLPSTLFCASVHEAVFEARQKERFEALFTIYDDQVTFQLFKSPRRVRINFSKP 97
Qy 62 LSAADARLRLHKTTEFLGKEMKLYFAQTLHG-----SSHLPAPPNDKQFLISPPASPPVGV 117
Db 98 EAAARARIELHETDFNGKQLKLYFAQVQMSGEVRDKSYLLPQPVKQFLISPPASPPVGV 157
Qy 118 KOVEDATPVNYDILYALSKLPGCEKVELHAADTPSPVVHVHVESDOENEEEMERMK 177
Db 158 KOSEDAMPVINYDILLCAVSKLPGCEKVELHAGTSTSPVVHVHVCSESETEEEEE-----T 211
Qy 178 KRPKPIIOTRRPE 191
Db 212 KNPKQKIAOTRRPD 225

```

## RESULT 8

```

ID_NLA DROME STANDARD; PRT; 292 AA.
AC Q9XZL8; Q9V391;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nebula protein.
GN NLA OR CG6072.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA McCormick A.V., Goldberg M.L.;
RA "Gene required for elongation of meiosis I spindle in Drosophila
RA females.";
RT Females.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Balleg R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotilier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fogle C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```





[illegible]

Db 158 LVHGLSDIVHG-LLELEAGLVGSSPTREEVEGCTEEVEGCTEEVEE 208

RESULT 13

NAH1\_CRIGR STANDARD; PRT; 822 AA.

AC PA8761;

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Sodium/hydrogen exchanger 1 (Na<sup>+</sup>/H<sup>+</sup>) exchanger 1 (NHE-1).

CN SLG9A1 OR NHE1.

OS Cricetulus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.

OC NCBI\_TaxID=10029;

OX NCBI\_TaxID=10029;

RP SEQUENCE FROM N. A.

RA MEDLINE=93192332; PubMed=8383540;

RA Councilon L., Pouyssegur J., "Nucleotide sequence of the Chinese hamster Na<sup>+</sup>/H<sup>+</sup> exchanger NHE1."

RL Biochim. Biophys. Acta 1172:343-345(1993).

CC -1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL TRANSDUCTION.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: PHOSPHORYLATED (POSSIBLE).

CC -1- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.

CC -1- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.

CC -1- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL, X68970; CAA48771.1; -

DR InterPro: IPR000676; NaH\_Exchange.

DR InterPro: IPR004709; NaH\_Exchange3.

DR Pfam: PF00999; Na\_H\_Exchange; 1.

DR PRINTS; PR01084; NAHEXCHNGR.

DR TIGRFAMs; TIGR00840; b\_gpal; 1.

KM Transmembrane; Glycoprotein; Transport; Antiport; Sodium transport; Multigene family; Phosphorylation.

KW

FT DOMAIN 1 12

FT DOMAIN 13 32

FT DOMAIN 105 127

FT DOMAIN 128 130

FT DOMAIN 131 150

FT DOMAIN 151 162

FT DOMAIN 163 183

FT DOMAIN 184 188

FT DOMAIN 189 210

FT DOMAIN 211 230

FT DOMAIN 231 251

FT DOMAIN 252 260

FT DOMAIN 261 282

FT DOMAIN 283 301

FT DOMAIN 302 322

FT DOMAIN 323 336

FT DOMAIN 337 357

FT DOMAIN 358 388

FT TRANSMEM 389 410

J (M8) (POTENTIAL).

FT DOMAIN 411 416

FT TRANSMEM 417 438

FT DOMAIN 439 452

FT DOMAIN 453 473

FT TRANSMEM 474 482

FT TRANSMEM 483 503

FT DOMAIN 504 822

FT CARBOHYD 374 374

SO SEQUENCE 822 AA; 92003 MW; E97C1ACD4EB8BDAA CRC64; (POTENTIAL).

Query Match 8.3%; Score 86.5; DB 1; Length 822;

Best Local Similarity 23.7%; Pred. No. 6.2; Indels 39; Gaps 8;

Matches 40; Conservative 26; Mismatches 64;

QY 44 TFQYKSFKEVRINFNSPLSA--DARLRHKTFLQKEMKLYFAQTLH----- 90

Db 646 TRQLRLSYNNHTL-VADPYEAWNMQLRQKARQLQKQNSVLTVPANHLDSPTMSRAR 704

QY 91 IGSSHLA-PNPDKQFLISPPASPPGKQVEDATPVINDLYAISKLAGERTELHAA 149

Db 705 IGSDPLAYPEKADLPVITIDPASP-----QSPSSVDLVNELKAKV-----LGVN 749

QY 150 TDPPTSVVHYCESDOENEBEEMERKRPKPKI-----IQTRRPETYP 194

Db 750 RDPY-----RLTRGEEDDEDGVMRKRPSSPGTDVFTAPMYSF 793

RESULT 14

GALE MYCPN STANDARD; PRT; 338 AA.

AC P75517;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-galactose 4-epimerase).

CN GALE OR MPN257 OR MP576.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI\_TaxID=2104;

RN [1]

RP SEQUENCE FROM N. A.

RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633;

RA Himmelfreisch R., Hilbert H., Plagene H., Pirkl E., Li B.-C., Herrmann R.;

RA "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."

RT Nucleic Acids Res. 24:4420-4449(1996).

RL

CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.

CC -1- COFACTOR: NAD.

CC -1- PATHWAY: Galactose metabolism; third step.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL, AE000056; AAB96224.1; -

DR HSSP; P09147; IYK5.

DR InterPro: IPR01509; Epimerase\_DH.

DR Pfam: PF01370; Epimerase; 1.

DR TIGRFAMs; TIGR01179; gale; 1.

KM Isomerase; NAD; Galactose metabolism; Complete proteome.

FT NP BIND 7 38

FT SEQUENCE 338 AA; 38132 MW; 9C50FP3856B8C03 CRC64;

Query Match 8.2%; Score 86; DB 1; Length 338;



**THIS PAGE BLANK (user)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 18:11:26 ; Search time 138.531 Seconds  
(without alignments)  
3202.482 Million cell updates/sec

Title: US-09-782-953-9  
Perfect score: 1041  
Sequence: 1 MFAP5MDCVSTLVACVVDV.....SPKPIQTRRFGLPSPSVN 197

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-Q=/cgn2\_1/USPTO spool/US0782953/runat\_11122002\_114428\_17330/app\_query.fasta\_1.1173  
-DB=N Geneseq 101002 -OPMT=fastap -SUFFIX=ring -MINMATCH=0 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human0.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0782953 @CGN 1 1 220 @runat\_11122002\_114428\_17330 -NCPU=6 -ICPU=3  
-NO XLPXY -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002:  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	100.0	594	24	AAD30154	Mouse MCIP associa
2	1007	96.7	3184	24	AAD30157	Human MCIP associa
3	1007	96.7	3184	24	ABL61768	Colon adenocarcino
4	1007	96.7	3184	24	ABL65187	Lung cancer relate
5	910	87.4	1021	24	ABA91458	Rat Down syndrome
6	905	86.9	934	22	AAF25338	Nucleotide sequenc
7	905	86.9	3159	24	ABA91457	Human Down syndrom
8	658.5	63.3	828	24	AAD30158	Human MCIP associa
9	641	61.6	2331	24	AAD30155	Human MCIP associa
10	633.5	60.9	720	24	AAD30159	Human MCIP associa
11	632.5	60.8	597	24	AAD30153	Mouse MCIP associa
12	628	60.3	2348	24	AAS94805	Human DNA sequence
13	626.5	60.2	2212	24	AAD30156	Human MCIP associa
14	613.5	58.9	597	24	AAD30152	Human MCIP associa
15	601	57.7	2173	24	ABN96843	Gene #3341 used to
16	600	57.6	2174	20	AX01282	Human DSCR1 coding
17	595	57.2	615	24	ABA91463	Rat Down syndrome
18	584.5	56.1	2358	21	AAF18328	Lung cancer associ
19	564	54.2	599	24	AAD30151	Human MCIP associa
20	404	38.8	1820	23	ABL29301	Drosophila melanog
21	404	38.8	12550	23	ABL29300	Drosophila melanog
22	323	31.0	442	21	AC01774	Human secreted pro
23	294	28.2	531	24	ABA91461	Rat Down syndrome
24	223	21.4	412	22	ABA56172	Human foetal liver
25	223	21.4	412	22	AAK04364	Human brain expres
26	223	21.4	412	22	AAI14449	Probe #4382 for ge
27	223	21.4	412	22	AAI35821	Probe #4507 used t
28	223	21.4	412	22	AAI04272	Probe #4263 used t
29	223	21.4	412	24	ABS04419	Human genome-deriv
30	223	21.4	446	22	ABA43406	Human breast cell
31	223	21.4	446	22	ABA53852	Human foetal liver
32	223	21.4	446	22	ABA23598	Probe #2064 for ge
33	223	21.4	446	22	AAK02113	Human brain expres
34	223	21.4	446	22	AAK27563	Human bone marrow
35	223	21.4	446	22	AAI12148	Probe #2081 for ge
36	223	21.4	446	22	AAI33493	Probe #2179 used t
37	223	21.4	446	22	AAI02061	Probe #2052 used t
38	223	21.4	446	24	ABS02041	Human genome-deriv
39	169	16.2	486	22	ABA42132	Human breast cell
40	169	16.2	486	22	ABA52554	Human foetal liver
41	169	16.2	486	22	ABA22343	Probe #809 for gen
42	169	16.2	486	22	AAK00816	Human brain expres
43	169	16.2	486	22	AAK36270	Human bone marrow
44	169	16.2	486	22	AAI10903	Probe #836 for gen
45	169	16.2	486	22	AAI32163	Probe #849 used to

ALIGNMENTS

RESULT 1  
AAD30154  
ID AAD30154 standard; DNA; 594 BP.  
XX  
AC AAD30154;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Mouse MCIP associated DNA #4.  
XX  
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
CDS 1..594  
FT /\*tag= a

```

TT      /product= "Mouse MCIP associated protein #4"
FX
FN      MO200204491-A2.
PD
XX      17-JAN-2002.
XX
XX      06-JUL-2001; 2001WO-US21662.
XX
XX      07-JUL-2000; 2000US-216601P.
PR      13-FEB-2001; 2001US-0782953.
XX
XX      (TEXA ) UNIV TEXAS SYSTEM.
XX      (WILL.) WILLIAMS S R.
XX      (ROTH/) ROTHERMEL B.
XX
XX      Williams SR, Rothermel B;
XX
XX      WPI; 2002-179698/23.
XX      P-PSDB; AAE18913.
XX
XX      Screening for modulators of muscle calcineurin interacting protein
XX      (MCIP) binding, expression or phosphorylation, useful for treating
XX      cardiac hypertrophy or heart failure, comprises mixing MCIP,
XX      calcineurin and a test compound -
XX
XX      Disclosure: Page 153-154; 174pp; English.
XX
XX      The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX      and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX      complex with the catalytic subunit of calcineurin and increased levels
XX      of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX      transcription of certain target genes. The invention also relates to
XX      methods for identifying modulators of MCIP binding, expression or
XX      phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX      may be used for treating cardiac hypertrophy and heart failure.
XX      Antibodies to MCIP can be used in characterizing the MCIP content of
XX      healthy and diseased tissues and subsequently for determining the
XX      presence or absence of cardiomyopathy or as predictor of heart disease.
XX      The present sequence is mouse MCIP associated DNA.
XX      Note: This sequence has been described as mouse MCIP1 in the
XX      specification, however the sequence seems to be a polynucleotide
XX      encoding a MCIP associated protein.
XX
XX      Sequence 594 BP; 148 A; 165 C; 149 G; 132 T; 0 other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.:          9.71e-99          Length:          594
XX      Score:              1041.00          Matches:          197
XX      Percent Similarity: 100.00%          Conservative:    0
XX      Best Local Similarity: 100.00%          Mismatches:     0
XX      Query Match:         100.00%          Gaps:           0
XX      DB:                  24
XX
XX      US-09-782-953-9 (1-197) x AAD30154 (1-594)
OY      1 MetPrCALAproSeMetAspCyGAAPVAlSeThrLeuValAlaCyValValAspVal 20
Db      1 AAGCGACGCCCTTAGCATGAGTATGTTTCCACTTCGGTCGCCCTGTGTGTGATGTC 60
OY      21 GIUValPheThrAsnGInGluValIyAsnGluYpPheGluGlyLeuPheArgThrTyrAsp 40
Db      61 GAGGCTTTTACCAATCAGAGAGGTTAAGGAAAATTCGAGGGACCTGTCGGGACCATCAT 120
OY      41 GIUCyValThrPheGlnLeuPheLysSerPheArgValAlaGlyLeuAsnPheSerHis 60
Db      121 GAATGTGTGACCTTCAGCTGTTTAAAGATTTCACACGGGTTCGATTAATTTACGCCAT 180
OY      61 ProLysSerAlaAlaArgValArgIleGluLeuHisAsnIuThrGlnPheArgGlyLysLys 80
Db      181 CCCAAATCTGCAGCCCGGTCGCCGATAGAGCTTCATGACATCAGTTTACAGAGGAAGAG 240
OY      81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100

```

Dd	241	CTVAAAGCTACTGCGCCGACGGTCCAGACCCCAGAGACAGATGGAGACAAGAAGTCATTGG	300
Oy	101	AlAAtProCgInPrOvAlAlYnGlnPheLeuIleSerProProSerSerProProVaIGly	120
Dd	301	GCACCTCCACAGAGCCTGGCAACAGTTCTTCATCTCACCCCCCTTCACTCTCTCTGTGGC	360
Oy	121	TrypSerProIleSerAspAlaThrProValIleuAsnTrpAspLeuLeuTyraIalaIa	140
Dd	361	TGGAGGCTTATCAGCGATGCCACACAGTCTCTAACATAAGACTTCTTAATGCTGTGGC	420
Oy	141	LysLeuGIlyProGlyGluLysTYrgIleuHisAlaGlyThnGlnSerThrProSerVal	160
Dd	421	AAACTAGGACCAGGAGAGAAATATAGAGCTTCACGCTCGAACTGAATGACTACACGAGCGTC	480
Oy	161	VaiValHisValCyaaPserAspMetGluGluGluGluAspProLyThrSerPolylys	180
Dd	481	GTTGGCATGTGTGTGACAGCGACATGGAGAGAGAGAGACCCAAAAGACTTCCCCCAAG	540
Oy	181	ProLyIleIleGlnThrArgArGProGlyLysProProSerValSeran	197
Dd	541	CCAAAAATCATTCAGACCCGAGCGCTGGGCTTGCCACCCCTCGGTCCAAC	591
<b>RESULT 2</b>			
ID	AAD30157	AAD30157 standard; DNA; 3184 BP.	
XX	AC	AAD30157;	
XX	DT	17-MAY-2002 (first entry)	
XX	DE	Human MCIP associated DNA #3.	
XX	KM	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;	
XX	OS	heart failure; cardiomyopathy; heart disease; human; gene; ds.	
XX	XX	Homo sapiens.	
XX	Key	Location/Qualifiers	
XX	FT	CDS	205..783
XX	FT	/*tag= a	/product= "Human MCIP associated protein #3"
XX	PN	W0200204491-A2.	
XX	PD	17-JAN-2002.	
XX	PF	06-JUL-2001; 2001MO-US21662.	
XX	PR	07-JUL-2000; 2000US-216601P.	
XX	PR	13-FEB-2001; 2001US-0782953.	
XX	PA	(TEXAS ) UNIV TEXAS SYSTEM.	
XX	PA	(WILL.) WILLIAMS S R.	
XX	PPA	(ROTH/) ROTHERMEL B.	
Pt	Williams SR,	Rothermel B;	
XX	DR	WPI; 2002-179698/23.	
XX	DR	P-PSDB; AAE18916.	
PT	Screening for modulators of muscle calcineurin interacting protein		
PT	(MCIP) binding, expression or phosphorylation, useful for treating		
PT	cardiac hypertrophy or heart failure, comprises mixing MCIP,		
PT	calcineurin and a test compound -		
XX	Disclosure; Page 163-165; 174pp; English.		
XX	The invention relates to muscle calcineurin interacting proteins (MCIPs)		
XX	and nucleic acid molecules encoding such proteins. MCIPs form a physical		
CC	complex with the catalytic subunit of calcineurin and increased levels		
CC	of MCIPs correspond to a reduced ability of calcineurin to stimulate		
CC	transcription of certain target genes. The invention also relates to		
CC	methods for identifying modulators of MCIP binding, expression or		

CC phosphorylation. Inhibitors or promoters of MCP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCP can be used in characterising the MCP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCP associated DNA.  
 CC Note: This sequence has been described as human MCP13 in the  
 CC specification, however the sequence seems to be a polynucleotide encoding  
 CC a MCP associated protein.

XX SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

## Alignment Scores:

Pred. No.: 2,88e-94 Length: 3184  
 Score: 1007.00 Matches: 190  
 Percent Similarity: 98.98% Conservatives: 5  
 Best Local Similarity: 96.45% Mismatches: 2  
 Query Match: 96.73% Indels: 0  
 DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x AAD30157 (1-3184)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
 Db 190 ATGCCAGCCCTAGCATGGAGCTGTGATGTTCCACTCTGGTTGCTGTGGTGATGTC 249  
 Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluLysLeuPheArgThrTyrAsp 40  
 Db 250 GAGGCTCTTACCAATCAGGAGGTTAAGGAAAAATTTGGGGGACTGTTTCGACTTATGAT 309  
 Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValAlaGlnPheSerHis 60  
 Db 310 GACTGTGTGAGCTTCCAGCTATTTAAGAGTTTTCAGAGCTGTCCGTATATAAATTCAGCAAT 369  
 Qy 61 ProllysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80  
 Db 370 CCTAAATCTGAGCCCGAGTAGATAGAGCTTCTGAACCCCAATTCAGAGGAAAAA 429  
 Qy 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
 Db 430 TTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGAGACAAATGCACCTG 489  
 Qy 101 AlaProGlnProAlaLysGlnPheLeuIleSerProSerSerProProValGly 120  
 Db 490 GCTCACCCCGAGCTGCAACAGTTTCTCATCTCGCCCTCTCTCCCTCCACCTGTTCAGC 549  
 Qy 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140  
 Db 550 TGGCAGCCCATCAACGATGCCAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCC 609  
 Qy 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160  
 Db 610 AAACAGGACGAGGAGAGAGATGAGCTCCATGACGGGACTGAGTCCACCCCAAGTGC 669  
 Qy 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180  
 Db 670 GTCTGTGACGTGTGCGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729  
 Qy 181 ProllysIleLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 730 CCAAAATATCCAAATCTGGGCTCTGGCTCCACCTCTCGTGTCCAAC 780

## RESULT 3

ABL61768  
 ID ABL61768 standard; DNA; 3184 BP.

XX ABL61768;

XX AC

DT 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:105.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX Homo sapiens.  
 OS  
 XX WO200194629-A2.  
 PN  
 XX 13-DEC-2001.  
 PD  
 XX  
 XX 30-MAY-2001; 2001WO-US10838.  
 PF  
 XX 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-231133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a  
 chemical agent to be tested for anti-neoplastic activity, and  
 determining a change in expression of a gene of a signature gene set -

Claim 1; SEQ ID 105; 44pp; English.

The present invention describes a method (M1) for screening for an  
 anti-neoplastic agent. The method involves exposing cells to a chemical  
 agent to be tested for anti-neoplastic activity, determining a change in  
 expression of at least one gene (I) of a signature gene set, where (I)  
 comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.

CC  
 XX SQ Sequence 3184 BP, 921 A; 681 C; 657 G; 925 T; 0 other;

# Alignment Scores:

Pred. No.:	2, 886-94	Length:	3184
Score:	1007.00	Matches:	190
Percent Similarity:	98.98%	Conservative:	5
Best Local Similarity:	96.45%	Mismatches:	2
Query Match:	96.73%	Indels:	0
DB:	24	Gaps:	0

US-09-782-953-9 (1-197) x ABL61768 (1-3184)

```

OY 1 MetProAlAProSerMetArCyAerValSerThrLeuValAlaCyAerValAlaAerVal 20
DB 190 ATGCCAGCCCTTACAGTGAAGTGTGATGTTTCCACTGCTGGTCCCTGGTGGAGTGC 249
OY 21 GluValAlaPheThrAngGlnGluValLysGluLysPheGlnGlyLeuPheArgThrTyrAsp 40
DB 250 GAGGTCTTTACCATAGGAGGTGTAAGAAATTGGGGGACTGTTGGACTTATGAT 309
OY 41 GluCyValAlaThrPheGlnLeuPheLysSerPheArgValAlaArgLysPheSerHis 60
DB 310 GACTGTGAGACGTTCCAGCTATTAGAGTTTGACAGTGTCCGTTTAACTTACCAAT 369
OY 61 ProLysSerAlaAlaArgAlaArgLysGluLeuHisGluThrGlnPheArgLysLys 80
DB 370 CCGAATTCGACGCCGAGCTAGATAGACTTCATGAAGAACCAATTCAGAGGAAAA 429
OY 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGlnThrAspGlnLysPheLysLeu 100
DB 430 TTAACCTCTACTTTCACAGGTTCAGACTCCAGACACAGATGAGACAAACTGCACTTG 489
OY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProPheSerProProValGly 120
DB 490 GCTCCACCCAGCCTGCCAAAGATTTCATCTCGCCCTTCTCCCACTGTAGC 549
OY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuLysTyrAlaValAla 140
DB 550 TGGCACCCTCATCAGATGCGCAGCTCTCACTATGAGCTCTCTATGCTGTGGCC 609
OY 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
DB 610 AAACTAGACACAGAGAGAAATGATGAGCTCCATGACGAGGACTAGTCCACCCCAATGTC 669
OY 161 ValValHisValLysAspSerAspMetGluLysGluLysAspProLysTyrThrSerProLys 180
DB 670 GTGTCACACCTGTGCACATGACATGAGAAAGAAAGAGAACCCAAAGATTTCCCAAAG 729
OY 181 ProLysGlnLeuGlnThrArgArgProGlyLeuProPheSerValSerAsn 197
DB 730 CCAGAAATCATCCAAACTGGCGCTCTGAGCTCCCAACCTCCGTGTCAAC 780

```

RESULT 4  
 ABL65187  
 ID ABL65187 standard; DNA; 3184 BP.  
 XX  
 AC ABL65187;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX

DE Lung cancer related gene sequence SEQ ID NO:3524.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
 KW gene; ds.

OS Homo sapiens.

XX MO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001KO-US10838.

XX 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-231133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.

XX (AVALON) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppe DR, Weaver Z;  
 XX

XX WPI, 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 XX  
 PS Claim 1; SEQ ID 3524; 44bp; English.

CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical



CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (1) of a signature gene set, where (1)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (1) has cytostatic  
CC an anti-neoplastic agent, and can be used in gene therapy. M1 can be used for screening  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX.

SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:  
Pred. No.: 2.88e-94 Length: 3184  
Score: 1007.00 Matches: 190  
Percent Similarity: 98.98% Conservative: 5  
Best Local Similarity: 96.45% Mismatches: 2  
Query Match: 96.73% Indels: 0  
DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x ABL65187 (1-3184)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
Db 190 ATGCAGCCCTAGCATGGAGCTGTGATGTTTCCACTCTGGTGTGCTGTGGATGTC 249  
Qy 21 GluValPheThrAenGlnGluValLysGluLysPheGluLysLeuPheArgThrTyrAsp 40  
Db 250 GAGGTCTTTACCAATCAGGAGGTTAAGCAAAATTTGGGGAGCTGTTTCGACTATGAT 309  
Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60  
Db 310 GACTGTGTGAGTTTCCAGCTATTAAAGAGTTTCAGAGCTGTCCGTATAAACTTCAGCAAT 369  
Qy 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80  
Db 370 CCTAAATCTGCAGCCCGAGCTAGATAGAGCTTCATGAAACCCCAATTCAGAGGAAATAA 429  
Qy 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
Db 430 TTAAGGCTCTACTTTGCACAGGTTCCAGCTCCAGACAGATGGAGCAAACTGCACCTG 489  
Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProSerSerProProValGly 120  
Db 490 GCTCCACCCCGAGCTGCCAAGAGTTTCTCATCTCGCCCTTCTCCTCCCACTGTAGC 549  
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140  
Db 550 TGGCAGCCCATCAACGATGCCAGCGAGTCTCACTATGACCTCTCTATGCTGTGGCC 609  
Qy 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160  
Db 610 AAATAGGACCCAGAGAGAGATGAGTCCATGCGGGAGCTCAGTCCACCCCAAGTGC 669  
Qy 161 ValValHisValCysAspSerAspMetGluGluGluAsnProLysThrSerProLys 180  
Db 670 GTGTGACGCTGTGCGACGTGACATAGAGAGAGAGAGAGACCAAGAGACTTCCCAAG 729  
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
Db 730 CCAAAATATCAAACTCGCGCTCTGGCTCCGACCTCCCTGCTGCTCAAC 780

RESULT 5  
ABA91458/c  
ID ABA91458 standard; cDNA; 1021 BP.  
XX  
AC ABA91458;

XX 18-APR-2002 (first entry)  
XX  
DE Rat Down syndrome critical region 1-like 1 protein related clone.  
XX  
XX Down syndrome critical region 1-like 1; DSCR1L1 alpha; rat;  
KW Down syndrome; Alzheimer's disease; dementia; transgenic animal;  
KW animal model; diagnosis; gene therapy; ss.  
OS Rattus norvegicus.  
XX  
XX WO200204513-A2.  
XX  
PD 17-JAN-2002.  
XX  
XX 11-JUL-2001; 2001WO-US21982.  
XX  
PR 11-JUL-2000; 2000US-0614474.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Loring JF, Tingley DW, Edwards CM, Streeter DG;  
XX  
DR WPI; 2002-164633/21.  
XX  
PT Novel Down syndrome critical region 1-like 1 protein and nucleic acid  
PT encoding the protein useful for diagnosis and treatment of Alzheimer's  
PT disease, Down syndrome and other forms of dementia -  
XX  
PS Claim 3; Page 48; 54pp; English.  
XX  
CC The present sequence is that of cDNA clone Incyte ID No: 219435-1,  
CC which shows 88% identity to nucleotides 443-954 of a cDNA clone  
CC encoding human Down syndrome critical region 1-like 1 protein  
CC (DSCR1L1 alpha, see ABA91457), and 87% identity to nucleotides  
CC 1040-1212 of the human sequence. The nucleic acid molecule is  
CC useful for producing transgenic cell lines or organisms which model  
CC human disorders and upon which potential therapeutic treatments  
CC for such disorders may be tested. Differential expression of the  
CC human DSCR1L1 alpha gene is diagnostic of Down syndrome, Alzheimer's  
CC disease and other forms of dementia.  
XX  
SQ Sequence 1021 BP; 237 A; 241 C; 286 G; 257 T; 0 other;

Alignment Scores:  
Pred. No.: 7.39e-85 Length: 1021  
Score: 910.00 Matches: 192  
Percent Similarity: 94.63% Conservative: 2  
Best Local Similarity: 93.66% Mismatches: 3  
Query Match: 87.42% Indels: 8  
DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x ABA91458 (1-1021)  
Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
Db 855 ATGCCAGCCCTAGCATGGAGCTGTGATGTTTCCACTCTGGTCCCTGTGTGGATGTC 796  
Qy 21 GluValPheThrAenGlnGluValLysGluLysPheGluLysLeuPheArgThrTyrAsp 40  
Db 795 GAGGTCTTTACCAATCAGGAGGTTAAGCAAAATTTGGGGAGCTGTTCCGGACCTATGAC 736  
Qy 41 GluCysValThrPheGlnPheLysSerPheArgArgValArg-IleAsnPheSerHis 60  
Db 735 GATGTGTGACGCTTTCAGCTGTTTAAAGAGTTTCCAGCGGTTCCGAAATAAATTCAGCCA 676  
Qy 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHisGlu-ThrGlnPheArg-GlyL 79  
Db 675 CCCCCAAGCTGCAGCCCGTGGTCCCGATAGGAGCTTCATGTGGACCCAGCTTCATCAGGGA 616  
Qy 79 ysLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGly-AspLysLeu 98  
Db 615 AGAGCTGAAACTTACTTTTCACAGGTCCAGACCCCGACAGACAGATGGTAGACAACTG 556

QY 99 His-LeuAlaProProGlnProAlaGlnPheLeuIleSerProProSerSerProPr 118  
 Db 555 CAGTTGGACCCCCCAACCGCTGCCAAGAGTTCTCTATCTACCCCTTATCTCTCC 496  
 QY 118 oValGlyTyrLysProIleSerAspAlaThrProValIleuAsn-TyrAspLeuLeuTyr 138  
 Db 495 COTTGGCTGGAGCCCTATCAGCGATGCCACACAGTCTCCACAGTACGACCTCTTATG 436  
 QY 138 lAvalAlaLysLeuGlyProGlyGlyLysTyrGlyLeuHisAlaGlyThrGlySerThrP 158  
 Db 435 CCGTGCCCAACTAGACGACGAGAGAAATATGAGCTGATCGGGAACCTGACTTACAC 376  
 QY 158 rGSeValValAlaHisValCysAspSerAspMetGluGluGluGluAspProLysThrS 178  
 Db 375 CAGAGCTTGTCTGTCACGTCTGTGACAGCACTTGAGAGAGAGAGATCCAAAGACTT 316  
 QY 178 exProLysProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 315 CCCCAGGCAAAATCATCCAGACCCGGCGTCTGCTGCTGACCTCCGTGTCACAC 257  
 RESULT 6  
 AAF25338  
 ID AAF25338 standard; cDNA; 934 BP.  
 AC AAF25338;  
 XX  
 DT 30-APR-2001 (first entry)  
 DE Nucleotide sequence of a human detoxification protein.  
 XX  
 KM Human; detoxification protein; DETX; cancer; leukaemia; melanoma;  
 KM adenocarcinoma; autoimmune disorder; inflammatory disorder;  
 KM rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
 KM psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
 KM actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key location/Qualifiers  
 FT CDS 20..724  
 FT /tag= a  
 FT /product= "detoxification protein"  
 FT sig\_peptide 20..100  
 FT /tag= b  
 XX  
 MO200104305-A2.  
 XX  
 PD 18-JAN-2001.  
 XX  
 PF 06-JUL-2000; 2000MO-US18509.  
 XX  
 PR 07-JUL-1999; 99US-0142678.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Tang YT, Yue H;  
 XX  
 DR MPI; 2001-147193/15.  
 DR P-PSDB; AAB11788.  
 XX  
 PT New human detoxification protein and polynucleotide, useful for  
 PT diagnosis, prevention and treatment of autoimmune/inflammatory  
 PT disorders and cell proliferative disorders including cancer -  
 XX  
 PS Claim 5; Page 79; 79pp; English.  
 XX  
 CC The present sequence encodes a human detoxification polypeptide (DETX).  
 CC DETX and its (ant)agonists are useful for preventing or treating  
 CC disorders associated with decreased or increased expression or activity  
 CC of DETX. DETX polypeptides are useful for screening compounds that  
 CC specifically binds to DETX and for identifying (ant)agonists.  
 CC Diseases prevented, treated and diagnosed include cancers (e.g.

CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,  
 CC breast, kidney, liver, pancreas, prostate and uterus),  
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,  
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative  
 CC colitis), bacterial, fungal, parasitic infections and cell  
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,  
 CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as  
 CC antagonists, as a targeting or delivery mechanism for bringing  
 CC pharmaceutical agents into contact with cells or tissues expressing  
 CC DETX and for diagnosis of DETX-related disorders.  
 XX  
 SO Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,176-84 Length: 934  
 Scores: 905.00 Matches: 171  
 Percent Similarity: 95.72% Conservative: 8  
 Best Local Similarity: 91.44% Mismatches: 8  
 Query Match: 86.94% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-782-953-9 (1-197) x AAF25338 (1-934)  
 QY 11 SerThrLeuValAlaCysValAlaAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 161 AACTGTTGTTGGTGGATGATGTCACCGACGTGTTGAGAGAGAGAGAGAGAGAGAG 220  
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
 Db 221 AAATTGAGGAGACTGTTTCGAGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAAGGT 280  
 QY 51 PheArgGValAlaGlyLeuAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
 Db 281 TTCAGACCTGTCGGATGATTAACCTTCAGCAATCTTAATTCGAGCCGCGATGAGATAGAG 340  
 QY 71 LeuHisGlnThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
 Db 341 CTTCATGAAACCCAAATTGAGAGGAGAAATTAAGCTCTACTTGGCAGAGTTCAAGCT 400  
 QY 91 ProGlnThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110  
 Db 401 CCAGAGACAGATGAGAGACAACTGCACTTGCTCCACCCGACCTGCCAAACAGTTCTC 460  
 QY 111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130  
 Db 461 ATCTGCCCCCTCTCTCCCACTGTTGCTGGCGCCCATCAAGATGCCACGCCCACTC 520  
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlyLysTyrGlyLeu 150  
 Db 521 CTCACCTATGACCTCTCTTATGCTGTGSCCAAACTAGACAGCAGAGAGATATAGCTC 580  
 QY 151 HisAlaGlyThrGlnSerThrProSerValValAlaHisValCysAspSerAspMetGlu 170  
 Db 581 CATGCAGGAGACTGATGCCACCCCAAGTGTGCGACAGTGTGCAAGATGACATAGAG 640  
 QY 171 GlnGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190  
 Db 641 GAAGAGAGAGACCAAGACTTCCCAAGCCAAATAATCATCAACTGCGGCTCTGGC 700  
 QY 191 LeuProProSerValSerAsn 197  
 Db 701 CTGCCACCTCCGTGTCACAC 721  
 RESULT 7  
 ABA91457  
 ID ABA91457 standard; cDNA; 3159 BP.  
 AC ABA91457;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX  
 DE Human Down syndrome critical region 1-like 1 protein cDNA.



PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

XX Example 1; Page 167-168; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC Note: This sequence has been described as splice variant of MCIP1  
 CC initiated by exon 4 in the specification, however the sequence seems  
 CC to be a polynucleotide encoding a MCIP associated protein.

XX Sequence 828 BP; 227 A; 188 C; 227 G; 186 T; 0 other;

#### Alignment Scores:

Pred. No.:	6,52e-59	Length:	828
Score:	658.50	Matches:	131
Percent Similarity:	79.14%	Conservative:	17
Best Local Similarity:	70.05%	Mismatches:	38
Query Match:	63.26%	Indels:	1
DB:	24	Gaps:	1

US-09-782-953-9 (1-197) x AAD30158 (1-828)

OY 11 SerThLeuValAlaCyValValaPvalGluValPheThrAnoInGluVallyGlu 30  
 DB 161 ACCTCAGCTTTTCTTGGACAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAA 220  
 OY 31 LysPheGluGlyLeuPheArgThrTyrAaPgluCyValValThrPheGlnLeuPheLysSer 50  
 DB 221 AGATTGGAACACCTTCCACCATCTTGATGACACAGCTTACCTTCAAGCTTTAAAC 280  
 OY 51 PheArgArgValAlaArgIleAenPheSerHisProlYsSerAlaAlaArgAlaArgIleGlu 70  
 DB 281 TTGAGAAGATCGAATTAATTTTCAGCAAACTGAAGCGGACAGAGAGAGAGAGAA 340  
 OY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
 DB 341 CTCACGAAACAGACTTCAATGGGAGAACTTAAGCTATATTTTGCACAGGTGCAGATG 400  
 OY 91 ProGluThrArgLysPheLysLeuHisLysLeuAlaProProlInProAlaLysGlnPheLeu 110  
 DB 401 TCCGCGCAAGTGGGGAGCAAGTCTATCTCTGCGCCGCCAGCTTGCAAGCAAGTTCCTC 460  
 OY 111 IleSerProProlSerSerProProValGlyTyrLysProlIleSerAparAlaThrProVal 130  
 DB 461 ATCTCCCTCCAGCCTCTCCCGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520  
 OY 131 LeuAsnTyrAapLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlyLysTyrGluLeu 150  
 DB 521 ATAAATATATATTAATTAATCTGCTGCTGCTTCCAAATTTGGACAGAGAGAGAAATATGA 580  
 OY 151 HisAlaGlyThrGluSerThrProSerValValValHisValCysAaPserAaPmerGlu 170  
 DB 581 CAGCGGGAACAGAGTGCACACCCAGCTGTGCTTATGCTCTGTAAGTGAACGAA 640  
 OY 171 GluGluGluAapProlYsThrSerProlYsProlYsIleGlnThrArgArgProGly 190  
 DB 641 GAGGAABAGAGCAAAA--AACCCCAACAGAAATTTGCCAGAGAGAGAGAGAGAGAG 697  
 OY 191 LeuProProlSerValSerAaPn 197

DB 698 CCTCCGACCGCAGCGTTGAAT 718

#### RESULT 9

AAD30155  
 ID AAD30155 standard; DNA; 2331 BP.

AC AAD30155;

DT 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #1.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

OS Homo sapiens.

FX Key Location/Qualifiers

FT CDS /tag= a /product= "Human MCIP associated protein #1"

PN MO200204491-A2.

PD 17-JAN-2002.

PP 06-JUL-2001; 2001WO-US21662.

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

PA (TEXA) UNIV TEXAS SYSTEM.

PA (WIL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

PI Williams SR, Rothermel B;

DR WPI; 2002-179698/23.

DR P-PSDB; AAE18914.

PS Claim 72; Page 155-157; 174pp; English.

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.

CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.

CC The present sequence is human MCIP associated DNA.

CC Note: This sequence has been described as a promoter in claim 72 of  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.

XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;

#### Alignment Scores:

Pred. No.:	1.63e-56	Length:	2331
Score:	641.00	Matches:	120
Percent Similarity:	79.68%	Conservative:	29
Best Local Similarity:	64.17%	Mismatches:	30
Query Match:	61.58%	Indels:	8
DB:	24	Gaps:	2

US-09-782-953-9 (1-197) x AAD30155 (1-2331)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 174 AGCTCCCTGATGCTGTGTCGCAACAGTGTATATCTTCAGCGAAGTGAACACAGGCC 233

QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
 Db 234 AAATTTGAGTCCCTCTTTAGGACGTATGACCAAGGACATCACCTTTTCAGTATTTTAAGAGC 293

QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
 Db 294 TTCAACGAGTCAGATAAATCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 353

QY 71 LeuHisGluThrGlnPheArgGlyLysLysLysLysLysLysLysLysLysLysLysLys 90  
 Db 354 CTGATAGAGTGTGTTCTGGGAAGGAATGAATGATATATATTTTGTCTGACACCTTACAC 413

QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110  
 Db 414 ATAGGAGCTCA-----CACCTGGCTCCGCCAATCCAGACAAAGCAGTTTCTG 461

QY 111 IleSerProProSerSerProProValGlyTyrPheLysProLysProLysProLysProLys 130  
 Db 462 ATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAGAGTGGAATGCGACCCCGAGTC 521

QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
 Db 522 ATAAACTATGATCTCTTATATGCCATCTCCAGCTGGCGGCGGGAAGATGATGAATTG 581

QY 151 HisAlaGlyThrGluSerThrProSerValValValHisValCysAspSerAspMetGlu 170  
 Db 582 CACGACGAGTGTACACCTCCCGCGGTGGTGTCTCATGTGTGTGAGGTGATCAAGAG 641

QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186  
 Db 642 AAGGAGGAGGAAGAGGAATGGAAGAAATGAGGAGACTTAAGCCAAATATATCCAGACC 701

QY 187 ArgArgProGlyLeuProPro 193  
 Db 702 AGGAGCGCGAGTACACGCCG 722

RESULT 10  
 ID AAD30159  
 AC AAD30159 standard; DNA; 720 BP.  
 AC AAD30159;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated DNA #5.  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 OS Homo sapiens.  
 FH Key  
 FT CDS  
 FT 2..640  
 FT /\*tag= a  
 FT /product= "Human MCIP associated protein #5"  
 XX WO200204491-A2.  
 XX 17-JAN-2002.  
 XX 06-JUL-2001; 2001WO-US21662.  
 XX 07-JUL-2000; 2000US-216601P.  
 XX 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 XX (WILL/) WILLIAMS S R.

(ROTH/) ROTHERMEL B.

Williams SR, Rothermel B;

MPI; 2002-179698/23.

P-PSDB; AAE18918.

Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

Disclosure; Page 170-171; 174pp; English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding, expression or phosphorylation may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterising the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is human MCIP associated DNA described in the invention.

Sequence 720 BP; 196 A; 168 C; 192 G; 164 T; 0 other;

## Alignment Scores:

Pred. No.: 2, 12e-56 Length: 720  
 Score: 633.50 Matches: 127  
 Percent Similarity: 77.01% Conservative: 17  
 Best Local Similarity: 67.91% Mismatches: 32  
 Query Match: 60.85% Indels: 11  
 DB: 24 Gaps: 2

US-09-782-953-9 (1-197) x AAD30159 (1-720)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 83 ACCTCACTTTTGTGTCGCGCTCCATGAAGCAGTGTGTAGGCACGAGAGGAGAA 142

QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
 Db 143 AGATTTGAGCACTCTTCCACATCTATGATGACCAAGTTACTTTTCAGCTGTAAAAAGC 202

QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
 Db 203 TTAGAAGAGTCAGATAAATTTTCAGCAACCTGAAGCGGCGAGAGCGGCAATAGAA 262

QY 71 LeuHisGluThrGlnPheArgGlyLysLysLysLysLysLysLysLysLysLysLysLys 90  
 Db 263 CTCCACGAAACAGACTTCAATGGCGAGAGCTAAAGCTATATTTTGCACAG----- 313

QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110  
 Db 314 -----TCCATCTCTGCGCGCCCGCTGTCAAGCAGTCTCTC 352

QY 111 IleSerProProSerSerProProValGlyTyrPheLysProLysProLysProLysProLys 130  
 Db 353 ATCTCCCTCCAGCCTCTCCCGAGTGGGTGGAAGCAGAGAGAGATGCCATGCTGT 412

QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
 Db 413 ATAAATATGATTTACTGTGCTGTTTCCAAATTTGGACCGAGAGAGAAATATCAACTT 472

QY 151 HisAlaGlyThrGluSerThrProSerValValValHisValCysAspSerAspMetGlu 170  
 Db 473 CACGCGGGAACAGAGTGCACACCGCTGGTGTTCATGCTGTGAAAGTGAACACTGAA 532

QY 171 GluGluGluAspProLysThrSerProLysIleIleGlnThrArgArgProGly 190



PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases  
 PT associated with atherosclerosis, comprises several polynucleotides that  
 PT are differentially expressed in foam cell development -

XX Claim 1; Page 112-113; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a  
 CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the  
 CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used  
 CC as PCR primers and probes. The polynucleotide sequences of the invention are also  
 CC useful in gene therapy. AAS94746-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation.

XX Sequence 2348 BP; 638 A; 473 C; 550 G; 587 T; 0 other;

## Alignment Scores:

Pred. No.: 3,66e-55 Length: 2348  
 Score: 628.00 Matches: 120  
 Percent Similarity: 79.26% Conservative: 29  
 Best Local Similarity: 63.83% Mismatches: 30  
 Query Match: 60.33% Indels: 9  
 DB: 24 Gaps: 2

US-09-782-953-9 (1-197) x AAS94805 (1-2348)

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 181 AGCTCCCTGATGCTGTGGCAGGATGATATCTTCAGCGAAAGTGAACACCGGGCC 240  
 Qy 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
 Db 241 AAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAGAC 300  
 Qy 51 PheArgValAlaGlnLeuAsnPheSerHisProLysSerAlaAlaArgAlaGlnGlu 70  
 Db 301 TTCAACGAGTCAGATTAATCTTACGAAACCTCTTCGCGACGATGCCAGGCTCCAG 360  
 Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
 Db 361 CTGATTAAGACTGAGTTCTGGGAGGAAATGAATATATTATTTGCTCAGACCTTACAC 420  
 Qy 91 ProGluThrAspGlyAspLysLeuHisLeu-AlaProProGlnProAlaLysGlnPheLe 110  
 Db 421 ATAGGAAGCTCA-----CACCTGGGCTCCGCCAAATCCAGACAGGAGTTTCT 468  
 Qy 110 uLeSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVa 130  
 Db 469 GATCTCCCTCCCGCCCTCTCCGCGAGTGGATGGAACAAGTGGAGATGCGACCCCACT 528  
 Qy 130 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlyLysTyrGluLe 150  
 Db 529 CATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGATATGAAT 588  
 Qy 150 uHisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetG1 170  
 Db 589 GCACGACGAGCTGACACCACTCCAGCGTGGTGGTCCATGTATGTCAGAGTATCAAGA 648  
 Qy 170 uGluGluGluAspProLysThrSer-----ProLysProLysIlelleGlnTh 186  
 Db 649 GAAGGAGGAAGAGAGGAAATGGAAGATGAGGAGACCTTACCAAGCAAAATATTCAGAC 708  
 Qy 186 rArgArgProGlyLeuProPro 193

Db 709 CAGGAGCGCGAGTACACGCCG 730

## RESULT 13

AAD30156  
 ID AAD30156 standard; DNA; 2212 BP.

XX AAD30156;

XX 17-MAY-2002 (first entry)

XX Human MCIP associated DNA #2.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 25..618

FT /\*tag= a

FT /product= "Human MCIP associated protein #2"

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

XX P-PSDB; AAE18915.

XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

XX Example 1; Page 159-161; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.

XX Antibodies to MCIP can be used in characterising the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.

XX The present sequence is human MCIP associated DNA.

XX Note: This sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.

XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

## Alignment Scores:

Pred. No.: 4.85e-55 Length: 2212  
 Score: 626.50 Matches: 121  
 Percent Similarity: 74.63% Conservative: 29  
 Best Local Similarity: 60.20% Mismatches: 34  
 Query Match: 60.18% Indels: 17  
 DB: 24 Gaps: 3

US-09-782-953-9 (1-197) x AAD30156 (1-2212)

QY 6 MetAapCYAspVal-----SerThrLeuValAlaCys 16  
 DB 13 ATTGACTCGAGATGAGAGGTGAGCTTGACAGACCTGCCAGCCACCTCCCTGT 72  
 QY 17 ValValAspValGluValPheThrAsnGlnGluValGluLeuPheGluLeuPhe 36  
 DB 73 CACCTGGACCCCGCGCTGTCTTCGAGACGGCTGTCCGCCCAATTTCAGTCCCTTT 132  
 QY 37 AsgThrTyAspGluCysValThrPheGlnLeuPheLeuSerPheArgArgValArg 56  
 DB 133 AGAGCTATGACAGACGACATCACTTCAGTATTTAAAGCTTCAACAGCTCAGAAAT 192  
 QY 57 AsnPheSerHisProLySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPhe 76  
 DB 193 AACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCACTGCATTAAGACTGAGTT 252  
 QY 77 ArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 96  
 DB 253 CTGGGAAAGAAATGAATGATTTTCTCAGACCTTACATAGAGCTCA----- 306  
 QY 97 LysLeuHisLeuAlaProGlnProAlaLeuGlnPheLeuIleSerProSerSer 116  
 DB 307 -----CACCTGGCTCCGCCAAATCCAGACAGAGCTTCTGATCTCCCTCCGCTCT 360  
 QY 117 ProProValGlyTyrPheLeuSerProIleSerAspAlaThrProValLeuLeuLeuLeu 136  
 DB 361 CCGCCAGTGGAGTGAACACAGAGTGAAGTGCACCCAGCTCAATACATATGATCTCTTA 420  
 QY 137 TyrAlaValAlaLeuLeuGlyProGlyGluLeuTyrGlyLeuHisAlaGlyThrGluSer 156  
 DB 421 TATGCATCTCCCAAGCTGGGCGCAAGGAAAGTATGAATTGACACCGAGCTGACACC 480  
 QY 157 ThrProSerValValValHisValCysAspSerAspMetGluGluGluLeuPheLeu 176  
 DB 481 ACTCCAGCGTGGTGGTCCATGTATGTAGAGATGATCAAGAGAGAGAGAGAGAGAGAG 540  
 QY 177 ThrSer-----ProLySerProLyIleIleGlnThrArgArgProGlyLeuPro 192  
 DB 541 ATGGAAAGATGAG 600  
 QY 193 Pro 193  
 DB 601 CCG 603

RESULT 14  
 AAD30152  
 ID AAD30152 standard; DNA; 597 BP.  
 AC AAD30152;  
 AC AAD30152;  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated DNA #2.  
 XX  
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..597  
 FT /tag= a  
 FT /product= "Human MCIP associated protein #2"  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX

PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 DR WPI; 2002-179698/23.  
 DR P-PSDB; AAE18911.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Claim 95; Page 147-148; 174pp; English.  
 XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 XX  
 SQ Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 1,96e-54 Length: 597  
 Score: 613.50 Matches: 119  
 Percent Similarity: 75.00% Conservative: 22  
 Best Local Similarity: 63.30% Mismatches: 38  
 Query Match: 58.93% Indels: 9  
 DB: 24 Gaps: 2

US-09-782-953-9 (1-197) x AAD30152 (1-597)

QY 11 SerThrLeuValAlaCysValAlaAspValGluValPheThrAsnGlnGluValGlu 30  
 DB 31 AGCCGACCATGCGCTGCGACCTGAGCCGCGGTTCGTGAGAGGCTGTGCGGCC 90  
 QY 31 LysPheGluGluLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheSer 50  
 DB 91 AAATTGAATCCCTCTTCAAGACATATGACAGACACACCTTCATATTATTAAGAC 150  
 QY 51 PheArgArgValArgIleAsnPheSerHisProLySerAlaAlaArgAlaArgIleGlu 70  
 DB 151 TTCAACGATGCGGATTAACCTTCAGCAACCTTATCTGAGCGGAGCGAGCTGCGG 210  
 QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLeuLeuLeuTyrPheAlaGlnValGlnThr 90  
 DB 211 CTGCACAAAGACCGAGTCTCGTGGGAAGGAATGAAGTGTATTTGCTCAGACTTACAC 270  
 QY 91 ProGluThrAspGlyAspLyLeuHisLeuAlaProGlnProAlaLeuGlnPheLeu 110  
 DB 271 ATAGAAAGTTCA-----CACCTGGCTCCGCCCAATCCGCAACACAGTTCTC 318  
 QY 111 IleSerProProSerSerProProValGlyTyrPheLeuSerAspAlaThrProVal 130  
 DB 319 ATCTCCCTCCGCGCTCTCTCCGTTGGCTGAAACAGTAGAAGATGCCACCCCGTC 378  
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLeuGlyProGlyGluTyrGluLeu 150  
 DB 379 ATAAATTACGATCTTTATATGCTTCCTCAAGCTGGGCGCGAGAGAGAGATATCAACTG 438  
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170



```

Db 439 CATCGACGACAGACCCACTCCCGAGTGTGTGTCACGCTGTGTGAGATGACCAAGAG 498
QY 171 GluGluGluAapProLys-----ThrSerProLysProLysIleIleGln 185
Db 499 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558
QY 186 ThrArgArgProGlyLeuProPro 193
Db 559 ACACGGAGACCGGAGTACACACCG 582

RESULT 15
ABN96843
ID ABN96843 standard; DNA; 2173 BP.
XX
AC ABN96843;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3341 used to diagnose liver cancer.
XX
KW Gene; liver cancer; db; hepatocellular carcinoma; hepatotrophic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
FN W0200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer.
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 3341; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotrophic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;

```

## Alignment Scores:

Pred. No.:	2,08e-52	Length:	2173
Score:	601.00	Matches:	115
Percent Similarity:	79.77%	Conservative:	23
Best Local Similarity:	66.47%	Mismatches:	27
Query Match:	57.73%	Indels:	8
DB:	24	Gaps:	2

```

US-09-782-953-9 (1-197) x ABN96843 (1-2173)
QY 25 AsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAspGluCysValThr 44
Db 43 ACAGAAATGGTGTATCCCAAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCC 102
QY 45 PheGlnLeuPheLysSerPheArgArgValArgIleAsnPheSerHisProLysSerAla 64
Db 103 TTTTCAGTATTTTAAGAGCTTCAACCGAGTCAGATAAACTTCAGCAACCCCTTCTCCGCA 162
QY 65 AlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLysLeuTyr 84
Db 163 GCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGAAAGGAAATGAAGTTATAT 222
QY 85 PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProGln 104
Db 223 TTGTCTCAGACCTTACACATAGGAAGCTCA-----CACCTGGCTCGCCAAAT 270
QY 105 ProAlaLysGlnPheLeuLysSerProProSerSerProValGlyTrpLysProIle 124
Db 271 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCCAGTGGGATGGAACAAGTG 330
QY 125 SerAspAlaThrProValLeuLeuThrAspLeuLeuTyrAlaValAlaLysLeuGlyPro 144
Db 331 GAAGATGCCAGCCCGAGTCATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA 390
QY 145 GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValHisVal 164
Db 391 GGGGAAAGTATGAATTGCCAGCAGCTGACACCCAGCTCCAGCGGTGGTGTCCATGTA 450
QY 165 CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys 180
Db 451 TGTGAGAGTGTATCAAGAGAAGGAGGAGGAGGAAGAGGAATCGAAAGATGAGGAGACCTTAG 510
QY 181 ProLysIleIleGlnThrArgArgProGlyLeuProPro 193
Db 511 CCAAAAAATTTATCCAGACCAGGAGCGCGGAGTACACGCCG 549

```

Search completed: December 14, 2002, 20:09:11  
Job time : 144.865 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.7083 Seconds  
(without alignments)  
1651.161 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQLPSATIAHLDP.....RPXPKIIQTRRPRYPIHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21:.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911	86.9	198	11 Q91WQ4	Q91WQ4 mus musculus
2	839.5	80.1	171	4 Q96R03	Q96R03 homo sapien
3	627.5	59.9	243	11 Q8VTP4	Q8VTP4 mus musculus
4	608.5	58.1	192	11 Q8VTP5	Q8VTP5 mus musculus
5	155	14.8	29	4 Q9H2A1	Q9H2A1 homo sapien
6	139.5	13.3	315	3 Q9P5S9	Q9P5S9 neurospora
7	114.5	10.9	249	3 Q9P4A1	Q9P4A1 cryptococcu
8	94	9.0	318	10 Q9SL96	Q9SL96 arabidopsis
9	94	9.0	318	10 Q8VZ43	Q8VZ43 arabidopsis
10	92	8.8	328	10 Q9LSD3	Q9LSD3 arabidopsis
11	92	8.8	922	4 Q9NTH6	Q9NTH6 homo sapien
12	92	8.8	1455	4 Q9UPV0	Q9UPV0 homo sapien
13	91.5	8.7	1274	10 Q9ZQK2	Q9ZQK2 arabidopsis
14	90.5	8.6	4025	4 Q8N13	Q8N13 homo sapien
15	88.5	8.4	680	17 Q8YV88	Q8YV88 pyrobaculum
16	87.5	8.3	1477	5 Q9VB52	Q9VB52 drosophila

17	87	8.3	267	11 Q95JH6	Q95JH6 mus musculus
18	86	8.2	731	4 Q9UIY5	Q9UIY5 homo sapien
19	84.5	8.1	719	11 Q61281	Q61281 mus musculus
20	84	8.0	751	5 Q8SX18	Q8SX18 drosophila
21	84	8.0	817	5 Q9VHB4	Q9VHB4 drosophila
22	83	7.9	1173	11 Q63624	Q63624 rattus norv
23	82.5	7.9	241	5 Q9N2W9	Q9N2W9 caenorhabdi
24	82.5	7.9	560	6 Q46411	Q46411 bos taurus
25	82.5	7.9	560	11 Q9DBG6	Q9DBG6 mus musculus
26	82.5	7.9	611	5 Q8WTE9	Q8WTE9 drosophila
27	82	7.8	697	10 Q9CA19	Q9CA19 arabidopsis
28	82	7.8	1128	11 Q88442	Q88442 mus musculus
29	82	7.8	1477	5 Q76931	Q76931 drosophila
30	81.5	7.8	528	17 Q39668	Q39668 archaeoglob
31	81.5	7.8	590	11 Q9CW63	Q9CW63 mus musculus
32	81.5	7.8	952	11 Q8R3C6	Q8R3C6 mus musculus
33	81	7.7	645	11 Q8VDM7	Q8VDM7 mus musculus
34	81	7.7	708	11 Q91YS3	Q91YS3 mus musculus
35	81	7.7	1099	4 Q9Y6N4	Q9Y6N4 homo sapien
36	81	7.7	1312	4 Q9NR59	Q9NR59 homo sapien
37	81	7.7	1343	4 Q9H7N4	Q9H7N4 homo sapien
38	80	7.6	681	4 Q9UR18	Q9UR18 homo sapien
39	80	7.6	727	4 Q9BQS9	Q9BQS9 homo sapien
40	80	7.6	817	5 Q9VXK5	Q9VXK5 drosophila
41	79.5	7.6	797	16 Q67651	Q67651 aquifex aeo
42	79.5	7.6	1020	10 Q9LP46	Q9LP46 arabidopsis
43	79.5	7.6	1136	5 Q9XX01	Q9XX01 caenorhabdi
44	79.5	7.6	1146	16 Q910B0	Q910B0 pseudomonas
45	79.5	7.6	1557	16 Q86560	Q86560 streptomyce

## ALIGNMENTS

## RESULT 1

Q91WQ4 ID Q91WQ4 PRELIMINARY; PRT; 198 AA.  
AC Q91WQ4;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Down syndrome critical region homolog 1 (human).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013551; AAH13551.1; --  
SQ SEQUENCE 198 AA; 22851 MW; F017C68F18ACC187 CRC64;

Query Match 86.9%; Score 911; DB 11; Length 198;  
Best Local Similarity 93.1%; Pred. No. 3.3e-82;  
Matches 175; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY	11	SATIACHLDRPVFVGLCRKAFESLFTYDKDTTQYFKSKRVIRINFSNPLSAADARLR	70
DB	11	SSLIACVANDVFESETRAKFESLFTYDKDTTQYFKSKRVIRINFSNPLSAADARLR	70
QY	71	LHKTEPLGKMKLYFAQTUHGSSHLAPNPDKQLISPPASPPVGVKQVEDATPVIND	130
DB	71	LHKTEPLGKMKLYFAQTUHGSSHLAPNPDKQLISPPASPPVGVKQVEDATPVIND	130
QY	131	LLYAIKSLGPGKGYELHAATDTPSVVHVCSDDNEMERMKRPKPKIIQTRRP	190
DB	131	LLYAIKSLGPGKGYELHAATDTPSVVHVCSDDNEMERMKRPKPKIIQTRRP	190
QY	191	EYTPHLS 198	
DB	191	EYTPHLS 198	

## RESULT 2

Q96R03 PRELIMINARY; PRT; 171 AA.  
 AC Q96R03; 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE Down syndrome critical region protein 1.  
 GN DSCR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,  
 Qiang B.;  
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF400429; AAK92478.1; --  
 SO SEQUENCE 171 AA; 19823 MW; 133001AEAEDEF0BD9 CRC64;

Query Match 80.1%; Score 839.5; DB 4; Length 171;  
 Best Local Similarity 95.3%; Pred. No. 3.2e-75;  
 Matches 161; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 30 AKFESLFRYDQDTTFQYKSFPRVIRINFSNPLSAADARLRHKTPELGKEMKLYFAOTL 89  
 DB 4 AKFESLFRYDQDTTFQYKSFPRVIRINFSNPLSAADARLRHKTPELGKEMKLYFAOTL 63  
 QY 90 HIGSSHLAPNPDQFLISPPSPVGMQVEDATPVINYLVAISKLGPEKYLHAA 149  
 DB 64 HIGSSHLAPNPDQFLISPPSPVGMQVEDATPVINYLVAISKLGPEKYLHAA 123  
 QY 150 TDPFBSVVHVCESDQENEESEEMERMKRPKPKIIQTRRPETPIHLS 198  
 DB 124 TDPFBSVVHVCESDQENEESEEMERMKRPKPKIIQTRRPETPIHLS 171

## RESULT 3

Q8V1P4 PRELIMINARY; PRT; 243 AA.  
 ID Q8V1P4; 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4 beta.  
 GN ZAKI-4 BETA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.;  
 RT "calcineurin inhibitory protein ZAKI-4."  
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AB061525; BAB71955.1; --  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; ttm; 1.  
 SO SEQUENCE 243 AA; 27332 MW; 2CDBQIA810291851 CRC64;

Query Match 59.9%; Score 627.5; DB 11; Length 243;  
 Best Local Similarity 62.4%; Pred. No. 4.8e-54;  
 Matches 121; Conservative 24; Mismatches 40; Indels 9; Gaps 2;

QY 5 DLQDLPATACHIDPRVFDGLCRAKFESLFRYDQDTTFQYKSFPRVIRINFSNPLSA 64  
 DB 51 DFLDLPNSLACVAVHGVFEESKEFGLFRYDQDTTFQYKSFPRVIRINFSNPLSA 110  
 QY 65 ADARLRHKTPELGKEMKLYFAOTLHIGSS---HLAPNPDQFLISPPSPVGMQVEDATPV 120

DB 111 ARARIELHETQFQKGLKLYFAQVOTPEYDQDKLHLAPPOPAKQFLISPPSPVGMQVEDATPV 170  
 QY 121 EDATPVINYLVAISKLGPEKYLHAAATDPPSVVVHVCESDQENEESEEMERMKRP 180  
 DB 171 SDATPVINYLVAISKLGPEKYLHAAATDPPSVVVHVCESDQENEESEEMERMKRP 225  
 QY 181 KPPIQTRRPETPIHLS 194  
 DB 226 KPPIQTRRPETPIHLS 239

## RESULT 4

Q8V1P5 PRELIMINARY; PRT; 192 AA.  
 ID Q8V1P5; 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4.  
 GN ZAKI-4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.;  
 RT "calcineurin inhibitory protein ZAKI-4."  
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AB061524; BAB71955.1; --  
 SO SEQUENCE 192 AA; 21540 MW; B2DB9819F436A8 CRC64;

Query Match 58.1%; Score 608.5; DB 11; Length 192;  
 Best Local Similarity 62.8%; Pred. No. 2.6e-52;  
 Matches 118; Conservative 22; Mismatches 39; Indels 9; Gaps 2;

QY 11 SATACHIDPRVFDGLCRAKFESLFRYDQDTTFQYKSFPRVIRINFSNPLSAADARLR 70  
 DB 6 STLVACVVDYEVFTNQEVKEFGLFRYDQDTTFQYKSFPRVIRINFSNPLSAADARLR 65  
 QY 71 LKHTPELGKEMKLYFAOTLHIGSS---HLAPNPDQFLISPPSPVGMQVEDATPV 126  
 DB 66 LKHTPELGKEMKLYFAOTLHIGSS---HLAPNPDQFLISPPSPVGMQVEDATPV 125  
 QY 127 INYLVAISKLGPEKYLHAAATDPPSVVVHVCESDQENEESEEMERMKRPKPKIIQ 186  
 DB 126 LKYLVLVAVAKLGPEKYLHAAATDPPSVVVHVCESDQENEESEEMERMKRPKPKIIQ 180  
 QY 187 TRRPETPIHLS 194  
 DB 181 TRRPETPIHLS 188

## RESULT 5

Q9H2A1 PRELIMINARY; PRT; 29 AA.  
 ID Q9H2A1; 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
 DE DSCR1 (Fragment).  
 GN DSCR1 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ernak G., Morgan T., Davies K.J.A.;  
 RT "Differential expression of adap78 (DSCR1) gene isoforms 1 and 2 in  
 human tissues, with particular reference to neurons."  
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF303449; AAG40774.1; --  
 PT NON\_TER 29 29

SQ SEQUENCE 29 AA; 3244 MW; D8CB7D1CD1521082 CRC64;

Query Match 14.8%; Score 155; DB 4; Length 29;  
Best Local Similarity 100.0%; Pred. No. 1.6e-08;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVDLQDLPSTATIACHLDPRVFDGLCR 29  
Db 1 MEEVDLQDLPSTATIACHLDPRVFDGLCR 29

## RESULT 6

ID Q9P5S9 PRELIMINARY; PRT; 315 AA.  
AC Q9P5S9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Nebula related protein.  
GN B5022.230.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariales; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; ALJ55932; CAB91442.1; -;  
SQ SEQUENCE 315 AA; 34194 MW; 76792987653C3AB7 CRC64;

Query Match 13.3%; Score 139.5; DB 3; Length 315;  
Best Local Similarity 40.2%; Pred. No. 1.2e-05;  
Matches 39; Conservative 13; Mismatches 38; Indels 7; Gaps 3;

QY 49 KSPKRVINFSNPLSAAARLRLHKTFLGKEMKLYPAQTLHIGSS----HLAPPNDKQ 104  
Db 101 KSFRIIVTFEQAIAVRSVMDGEALGRCRVFGQTPIDVSAADKHLALPDAGKL 160  
QY 105 FLISPPASPPGVKQ-VEDA--TPVINYDLYAISKL 138  
Db 161 FFISSPPSPHDEQRMEDAPNTMVAEDLAELAKL 197

## RESULT 7

ID Q9P4A1 PRELIMINARY; PRT; 249 AA.  
AC Q9P4A1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Calcineurin-binding protein.  
GN CBP1.  
OS Cryptococcus neoformans var. neoformans.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
OX NCBI\_TaxID=40410;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H99;  
RX MEDLINE=20359261; PubMed=10899116;  
RA Gorlach J., Fox D.S., Cutler N.S., Cox G.M., Perfect J.R., Heitman J.;  
RT Identification and characterization of a highly conserved calcineurin  
binding protein, CBP1/calcipressin, in Cryptococcus neoformans.;  
RL EMBO J. 19:3618-3629(2000).  
DR EMBL; AF230799; AAF63734.1; -;  
SQ SEQUENCE 249 AA; 27293 MW; 7B757183FA5EBF09 CRC64;

Query Match 10.9%; Score 114.5; DB 3; Length 249;  
Best Local Similarity 27.9%; Pred. No. 0.0027;  
Matches 46; Conservative 23; Mismatches 61; Indels 35; Gaps 8;

QY 10 PSATIACHL-DRPVF---VDGLCRAPESLRTYDKDTTFQYFKSKRVINFSNPLSAA 65  
Db 18 PTNTLALLPHPTLFAPPVLDLLRAHYEH---FGRIVHWAPVGRGAIWVFSEEEAE 73  
QY 66 DAR-----LRL-----HKTEFLGKEMKLYPAQTLHI-----GSSHLAPPNDK 103  
Db 74 NAKRQGDWLKLDVPVGEEKVONEGKLDIELVLRHLHLPPTTLNPDPAATHLAPPPLPH 133

QY 104 QFLISPPASPPGVKQVEDATP---VINYDLYAIS--KLGPCK 143  
Db 134 NFLISPPGSPGEGWEPAAEAPNRILPEDLQRALETLELNSGSK 178

## RESULT 8

ID Q9SL96 PRELIMINARY; PRT; 318 AA.  
AC Q9SL96;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE At2G25670 protein.  
GN AT2G25670.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosidia II; Brassicales; Brassicaceae; Arabidopsia.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
RA Buel C.R., Ketchum K.A., Lee J.J., Ranning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., VanAken S.E., Umavay L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrara A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana.";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006053; AAD31367.1; -;  
SQ SEQUENCE 318 AA; 35128 MW; 2F2E2CD269FFD004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;  
Best Local Similarity 25.2%; Pred. No. 0.39;  
Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

QY 54 VRINFSNPLSAAARLRLHKTFLGKEMKLYPAQTLHIGSSHLAPP--NPDKQFLISPPA 111  
Db 13 ITIQSTNLPALDTTKKKKSKAGSK-----GSSKREPEKEPEQVYVWATP 62  
QY 112 SPVGVKQVEDATPVINYDLYAISKLGPCKYELHAATDTPPSVVHVHVCESDQ----- 166  
Db 63 LKVKSWADIDDDDDDDYYATTAPPQSGWSTLPSH-----TDSKDVHVESESEDILD 117  
QY 167 -----NEEEHEEMRMKRPKPIQTRPE 191  
Db 118 EGDDVVEEQEQETEVOVHPPEV--KKAPE 146

## RESULT 9

Q8VZ43 PRELIMINARY; PRT; 318 AA.  
ID Q8VZ43

AC Q8V243;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 35.2 kDa protein.  
 GN ATG25670.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,  
 Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Mu H.C.,  
 Yamamura Y., Yu G., Xu S., Bowser L., Carninci P., Chen H., Chou R.,  
 Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,  
 Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,  
 Narusaka M., Nguyen M., Palm C.J., Sakurai T., Saitou M., Seki M.,  
 Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,  
 Theologis A.,  
 RT "Full length cDNA of gene At2g25670 (GI:15225169)."  
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY065267; AAL38743.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 318 AA; 35154 MW; 2F2E2CC278FED004 CRC64;  
 Query Match 9.0%; Score 94; DB 10; Length 318;  
 Best Local Similarity 25.2%; Pred. No. 0.39; Indels 30; Gaps 5;  
 Matches 38; Conservative 19; Mismatches 64;  
 QY 54 VRINFSNPADARLRLHTEFLGKEMKLYFAQTLLHGSSHLAPP--NPDKQFLISPPA 111  
 DB 13 ITIGSTNLPAALDRKKKKSKDKAGSK-----GSKSRREKPPQVVAAPRP 62  
 QY 112 SPVGVKQVEDATPVINYDLVAISLKGGEKYLHAATDPFPPSVVAVHCESDQF----- 166  
 DB 63 LKVKSWADIDDEDDDDVATTAPOSQWSTLPSH-----TDSKDVHVESESEEDILD 117  
 QY 167 -----NEEEEMERMRKPKPKIOTRRPE 191  
 DB 118 EGDVVEEEOETEVOHPEPV--KKAPE 146  
 RESULT 10  
 Q9LSD3 PRELIMINARY; PRT; 328 AA.  
 AC Q9LSD3;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Genomic DNA, chromosome 3, pl clone: MOJ10.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamitsu E., Tabata S.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 GN [12]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=COLUMBIA;  
 RX MEDLINE=20277480; PubMed=10819329;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 clones."  
 RT DNA Rep. 7:131-135(2000).  
 DR EMBL: AB026649; BAB01084.1;  
 DR InterPro: IPR001776; Aetolysin.

DR PROSITE; PS00274; AETOLYSIN; UNKNOWN.1.  
 SQ SEQUENCE 328 AA; 37438 MW; A43A5FEIAD0634AA CRC64;  
 Query Match 8.8%; Score 92; DB 10; Length 328;  
 Best Local Similarity 21.4%; Pred. No. 0.64; Indels 84; Gaps 9;  
 Matches 50; Conservative 24; Mismatches 76;  
 QY 23 FVDGLCRAKESLERTYKDTTFQYKSPKRVIN-PS-----NPLSADARLRLH 73  
 DB 31 FKGLICSLR-----AQVFSDVDIRTNSTFSFGTHDPNP-SKVQDLRFDE 76  
 QY 74 TFLGKEMKLYFAQTLLHGSSHLAPPD-----KQFLIS 108  
 DB 77 DEFG-----FLAIGLGTDPETPKFSAMVAEEDATGEIKEMAKIYAKLDOFLKE 127  
 QY 109 PPAPSPVGVKQVEDATPVINYDL-----LYAISKLGPEKY 144  
 DB 128 YPEDTRSKRVKISNECLQDYDLFRSIELTQSGNGRVKKKSLTSLFKRRQTVGGEPI 187  
 QY 145 -ELHAATDPFPPSVV--VHVCESDQNEEEEMERMRKPKPKIOTRRPEYTP 195  
 DB 188 IEKSTDAIKRVFKLHGASSTKRNDEDDSMKSKKDKLVQTCRRKHPV 241  
 RESULT 11  
 Q9NTH6 PRELIMINARY; PRT; 992 AA.  
 AC Q9NTH6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE Hypothetical 112.6 kDa protein (fragment).  
 GN DKZP43402413.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=TESTIS;  
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL137265; CAB70664.1;  
 KW Hypothetical protein.  
 FT NON-TER  
 SQ SEQUENCE 992 AA; 112628 MW; 039DF5B1B97E02F CRC64;  
 Query Match 8.8%; Score 92; DB 4; Length 992;  
 Best Local Similarity 22.7%; Pred. No. 2.8;  
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;  
 QY 96 LAPNPDKQFLISPPAS-----PPVG-----WKQVEDATPVINYDLVAIS----- 136  
 DB 245 LSPPLPHBERAQSPPRSILATEEPPQGEPOPEWEKAE-----LGEPSAASLSLQSLQR 300  
 QY 137 -----KLGRG-----EKYELHAATDPFPPSV----- 156  
 DB 301 EQAPSPAPACEKGKQSHQAEELGPGQEADEBEKVAVSPFPPSPVETKSTEPVAPPE 360  
 QY 157 -----VHVCESDQNEEEEMERMRKPKPKIOTRRPEYTP 196  
 DB 361 LSEALKAMEENVAQVLEDDQKHLLSKQ-EKMQQLRREKLQEEBEBELRLH 411  
 RESULT 12  
 Q9UPV0 PRELIMINARY; PRT; 1455 AA.  
 AC Q9UPV0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 GN KIAA1052 protein.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=99397452; PubMed=10470851;  
 RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,  
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 6:197-205(1999).  
 DR EMBL; AB028975; BAA83004.1; -  
 DR InterPro; IPR001202; WW\_Rep5\_WWP.  
 DR Pfam; PF00397; WW; 1.  
 DR SMART; SM00456; WW; 1.  
 DR PROSITE; PS50020; WW DOMAIN 2; 1.  
 SQ SEQUENCE 1455 AA; 163543 MW; 7F48093100C34819 CRC64;  
 Query Match 8.8%; Score 92; DB 4; Length 1455;  
 Best Local Similarity 22.7%; Pred. No. 4.5;  
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;  
 QY 96 LAPNPDQFLISPPAS-----PPVG-----WKVEDATPVINYLLVAIS----- 136  
 DB 474 LSPPLPHEAQSPPRSLATPEEPQPEQPEWKEAE-----LGEDSAASLSLSLSLQR 529  
 QY 137 -----KLGPG---EKYELHAATDTPSV----- 156  
 DB 530 EQAPSPPAECGKEQSQAEELGPGQBEADPEEKVAVSPTPVSVSEVSTEPVAPPEQ 589  
 QY 157 -----VVHVCSDQENEEEMERMKRPKPIIQTRRPYTPH 196  
 DB 590 LSEALKAMEAVQVLEQDQRLHLSKQ-EKMQQLREKLQEEEEEILRLH 640  
 RESULT 13  
 Q9ZQK2 PRELIMINARY; PRT; 1274 AA.  
 ID Q9ZQK2  
 AC Q9ZQK2  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Putative retroelement pol polyprotein.  
 GN AT2G07400.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffatt K.S.,  
 RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Smerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX Lin X.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006217; AAD1532.1; -  
 DR InterPro; IPR004312; ATHILA.

DR InterPro; IPR003653; SUMO\_protease.  
 DR Pfam; PF03078; ATHILA\_1  
 DR Pfam; PF02902; Peptidase\_C48; 1.  
 KW Polyprotein.  
 SQ SEQUENCE 1274 AA; 145232 MW; A2342B5FDCE45631 CRC64;  
 Query Match 8.7%; Score 91.5; DB 10; Length 1274;  
 Best Local Similarity 23.7%; Pred. No. 4.3;  
 Matches 51; Conservative 30; Mismatches 83; Indels 51; Gaps 9;  
 QY 24 VDGLCRKAFSLRPTVDKDTTFQYFKSKRVRINFNSPLSAADAR-LRLHKTEFLGKE-- 80  
 DB 673 IESRVEAKFEGFSGIENDV-----KQIKQLKAIADSKSSVIRDMYLAKTOPQTDQN 727  
 QY 81 MKLYFAQTLLH-----GSSHLAPPND-----KQFLISPPA-----SPPVGWK 118  
 DB 728 PKVQTQOTPDVPKPTNNQFATSPSPSKQADVKKKEKTINELIQPPGRRGKPSQPK 787  
 QY 119 QVEDATPVINYLLVAISKLGPGKVELHA-----ATDTPSVVVHVCSDQENEEEEE 173  
 DB 788 KVSPTWPKI-----TIKKLPQSEAKAEEDSDVVVDKVVSEYNEMLPSEDEDEEE 841  
 QY 174 MERMKRPKPIIOTRR-----PEYTPHLS 198  
 DB 842 TERLKSVMKIRLKTAVKLAPDGRSLINPEAMPHTHS 876  
 RESULT 14  
 Q9NR13 PRELIMINARY; PRT; 4025 AA.  
 ID Q9NR13  
 AC Q9NR13  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ALR-like protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CERVICAL CARCINOMA;  
 RX Chow V.T.K., Tan Y.C.;  
 RA "ALR-like protein, a novel human cDNA whose product is homologous to  
 RA the ALR protein.";  
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.  
 DR EMBL; AF264750; AAF74766.2; -  
 DR InterPro; IPR000194; ATPase\_a/bcentre.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR InterPro; IPR002219; DAG\_P8-bind.  
 DR InterPro; IPR003889; FYIich\_C.  
 DR InterPro; IPR003888; FYrich\_N.  
 DR InterPro; IPR000910; HMG\_12\_box.  
 DR InterPro; IPR003616; PostSET.  
 DR InterPro; IPR001214; SET.  
 DR InterPro; IPR001965; Znf\_PHD.  
 DR Pfam; PF00505; HMG\_box; 1.  
 DR Pfam; PF00628; PHD; 3.  
 DR Pfam; PF00856; SET; 1.  
 DR SMART; SM00109; C1; 1.  
 DR SMART; SM00542; FYRC; 1.  
 DR SMART; SM00541; FYRN; 1.  
 DR SMART; SM00398; HMG; 1.  
 DR SMART; SM00249; PHD; 4.  
 DR SMART; SM00508; PostSET; 1.  
 DR SMART; SM00184; RING; 3.  
 DR SMART; SM00317; SET; 1.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; UNKNOWN\_1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE; PS50280; SET; 1.  
 SQ SEQUENCE 4025 AA; 443532 MW; 41624149C28E4BDE CRC64;

Query Match 8.6%, Score 50.5; DB 4; Length 4025;  
 Best Local Similarity 27.9%, Pred. No. 24;  
 Matches 31; Conservative 15; Mismatches 46; Indels 19; Gaps 4;

QY 87 QTLHGSHLAPNDKOP--LISPPAPVPG--MKVQEDATPVINIDLYAISKLGPG 142  
 DB 941 QVSPGSSNSRPPSPMDPYAKWGTGPRPPVGHSPSRNSAPVENCPTPLSSVSR----- 995  
 QY 143 KYELHAATDPPTPSVYVHVCSQDENEEEMERKRPKXIOTRPEYT 193  
 DB 996 PLQMKETTANRPSVPRDLCSSTTND-----PYAKPPDTPRPMT 1036

## RESULT 15

Q82Y88 PRELIMINARY; PRT; 680 AA.  
 ID Q82Y88  
 AC Q82Y88;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE DNA replication licensing factor (lcm).  
 GN PAE901.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 CX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 aerophilum";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 DR EMBL; AE009792; AAL63108.1; -.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR001208; MCM.  
 DR Pfam; PF00493; MCM; 1.  
 DR ProDom; PD001041; MCM; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00350; MCM; 1.  
 DR PROSITE; PS50051; MCM\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 680 AA; 76173 MW; 79409981D729D151 CRC64;

Query Match 8.4%, Score 88.5; DB 17; Length 680;  
 Best Local Similarity 23.6%, Pred. No. 3.7;  
 Matches 50; Conservative 30; Mismatches 59; Indels 73; Gaps 12;

QY 40 DKDTTQYKSPKRVINFSNPLSADARLRHLKTEFLGKMKL----- 83  
 DB 80 DEPTAKALKRFYFRV--GSPLSVS---LRKLSEYIGRLIKIEGIVTRQTPPKHFLYK 133  
 QY 84 -----YFAOTLHGSHLAP-----NPDQFLISPPASPPVGMKV-----ED 122  
 DB 134 ALYRCQCGEIEILLOELHVEPRAKPCRCGASKSFILVTELSQYIDWKIVGERPD 193  
 QY 123 ATP-----VINVDLVAISKLGGEKVELHAATDPT-----PSVY-----VHV 160  
 DB 194 LPPGQLPRSVVVLDDL--VDTVKPGDIIISLGVVDLTLSLKKGRPPIVTSYIGVHV 251  
 QY 161 CESDQENEEEMERKRPKXIOTRPE 191  
 DB 252 ---DTMKKELVEI--TKEDQKILEISRPD 278

Search completed: December 11, 2002, 11:41:02  
 Job time : 28.7063 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.7184 Seconds  
(without alignments)  
858.887 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEVDQLDLPSTAIACHLDP.....RPKPIQTRPEYTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 101002:\*

- 1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1048	100.0	198	23 AAE18911 Human MCIP associa
2	1002.5	95.7	197	23 AAE18915 Human MCIP associa
3	920.5	87.8	198	23 AAE18912 Mouse MCIP associa
4	872.5	83.3	197	23 AAE18914 Human MCIP associa
5	835.5	79.7	171	20 AA073898 Human Down's Syndr
6	620.5	59.2	234	20 AAB31788 Amino acid sequenc
7	620.5	59.2	255	23 AAMS0760 Human Down syndrom
8	613.5	58.5	197	23 AAE18913 Mouse MCIP associa
9	596.5	56.9	192	23 AAE18916 Human MCIP associa
10	596	56.9	241	23 AAE18917 Human MCIP associa

11	595	56.8	212	23 AAE18918 Human MCIP associa
12	497	47.4	142	21 AAB58452 Lung cancer associ
13	439	41.9	111	21 AAG01768 Human secreted pro
14	379.5	36.2	292	22 ABB71467 Drosophila melanog
15	284	27.1	58	22 ABB29495 Peptide #2146 enco
16	284	27.1	58	22 ABB34672 Peptide #2178 enco
17	284	27.1	58	22 ABB36877 Peptide #4383 enco
18	284	27.1	58	22 ABB20081 Protein #2080 enco
19	284	27.1	58	22 AAM55464 Human brain expres
20	284	27.1	58	22 AAM57623 Human brain expres
21	284	27.1	58	22 AAM67851 Human bone marrow
22	284	27.1	58	22 AAM15671 Peptide #2105 enco
23	284	27.1	58	22 AAM17859 Peptide #4293 enco
24	284	27.1	58	22 AAM28174 Peptide #2211 enco
25	284	27.1	58	22 AAM30365 Peptide #4402 enco
26	284	27.1	58	22 AAM03406 Peptide #2088 enco
27	284	27.1	58	22 AAM05506 Peptide #4188 enco
28	284	27.1	58	23 AAB37397 Human peptide enco
29	284	27.1	58	23 AAB39657 Human peptide enco
30	271.5	25.9	56	22 ABB28178 Human peptide #829
31	271.5	25.9	56	22 ABB33353 Peptide #859 enco
32	271.5	25.9	56	22 ABB18813 Protein #812 enco
33	271.5	25.9	56	22 AAM54139 Human brain expres
34	271.5	25.9	56	22 AAM66532 Human bone marrow
35	271.5	25.9	56	22 AAM14405 Peptide #839 enco
36	271.5	25.9	56	22 AAM26818 Peptide #855 enco
37	271.5	25.9	56	22 AAM02132 Peptide #814 enco
38	271.5	25.9	56	23 AAB36184 Human peptide enco
39	98	9.4	19	23 AAE18921 Arabidopsis thalia
40	94	9.0	318	21 AAG30396 Arabidopsis thalia
41	94	9.0	318	21 AAG44378 Arabidopsis thalia
42	94	9.0	318	21 AAG49385 Arabidopsis thalia
43	94	9.0	318	21 AAG49386 Arabidopsis thalia
44	92	8.8	954	22 AAU14615 Novel bone marrow
45	90.5	8.6	4019	22 AAE13839 Human lung tumour-

#### ALIGNMENTS

RESULT 1

AAE18911

ID AAE18911 standard; Protein; 198 AA.

XX AAE18911;

AC AAE18911;

XX

DT 17-MAY-2002 (first entry)

XX

DE Human MCIP associated protein #2.

XX

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; human.

XX

OS Homo sapiens.

XX

PN WO200204491-A2.

XX

PD 17-JAN-2002.

XX

PF 06-JUL-2001; 2001WO-US21662.

XX

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX

PA (TEXA ) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

DR N-PSDB; AAD30152.

XX

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure: Page 148-149; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC  
 SQ Sequence 198 AA:  
 Query Match 100.0%; Score 1048; DB 23; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-108; Indels 0; Gaps 0;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAPESLFRITDKDTTFQYFKSKRVIRINFSN 60  
 DB 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAPESLFRITDKDTTFQYFKSKRVIRINFSN 60  
 QY 61 PLSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 DB 61 PLSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 QY 121 EDATPVINVDLLVAISKLGGEKVELHAATDTPSVVVHVCSDENEEEMERMRKP 180  
 DB 121 EDATPVINVDLLVAISKLGGEKVELHAATDTPSVVVHVCSDENEEEMERMRKP 180  
 QY 181 KPRTIOTRRREPTPIHLS 198  
 DB 181 KPRTIOTRRREPTPIHLS 198  
 RESULT 2  
 AAE18915  
 ID AAE18915 standard; Protein; 197 AA.  
 AC AAE18915;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #2.  
 XX  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 XX  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX

DR WPI; 2002-179698/23.  
 DR N-PSDB; AAD30156.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Example 1; Page 161-162; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 SQ Sequence 197 AA:  
 Query Match 95.7%; Score 1002.5; DB 23; Length 197;  
 Best Local Similarity 96.5%; Pred. No. 1.9e-103; Indels 1; Gaps 1;  
 Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAPESLFRITDKDTTFQYFKSKRVIRINFSN 60  
 DB 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAPESLFRITDKDTTFQYFKSKRVIRINFSN 60  
 QY 61 PLSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 DB 61 PLSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 QY 121 EDATPVINVDLLVAISKLGGEKVELHAATDTPSVVVHVCSDENEEEMERMRKP 180  
 DB 121 EDATPVINVDLLVAISKLGGEKVELHAATDTPSVVVHVCSDENEEEMERMRKP 179  
 QY 181 KPRTIOTRRREPTPIHLS 198  
 DB 180 KPRTIOTRRREPTPIHLS 197  
 RESULT 3  
 AAE18912  
 ID AAE18912 standard; Protein; 198 AA.  
 AC AAE18912;  
 DT 17-MAY-2002 (first entry)  
 DE Mouse MCIP associated protein #3.  
 XX  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; mouse.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 XX  
 PR 13-FEB-2001; 2001US-0782953.  
 XX

PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 XX Williams SR, Rothermel B;  
 XX  
 DR WPI: 2002-179698/23.  
 DR N-PSDB; AAD30153.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 XX (MCIP) binding, expression or phosphorylation, useful for treating  
 XX cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 XX calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 151-152; 174pp; English.  
 XX  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 XX and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 XX complex with the catalytic subunit of calcineurin and increased levels  
 XX of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 XX transcription of certain target genes. The invention also relates to  
 XX methods for identifying modulators of MCIP binding, expression or  
 XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 XX may be used for treating cardiac hypertrophy and heart failure.  
 XX Antibodies to MCIP can be used in characterising the MCIP content of  
 XX healthy and diseased tissues and subsequently for determining the  
 XX presence or absence of cardiomyopathy or as predictor of heart disease.  
 XX The present sequence is mouse MCIP associated protein.  
 XX Note: This sequence has been described as human MCIP3 in the  
 XX specification, however the sequence seems to be a MCIP associated  
 XX protein.  
 XX  
 SQ Sequence 198 AA;  
 Query Match 87.8%; Score 920.5; DB 23; Length 198;  
 Best Local Similarity 89.9%; Pred. No. 2.8e-94;  
 Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;  
 QY 4 VDLQDLP--SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSRVRINFN 60  
 DB 1 MDRFDSYNSFLIACVANDDFSESETRAKFESLFRITYDKDTTFQYFKSRVRINFN 60  
 QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120  
 QY 121 EDATPVINDLLVAISKLGPGKGYELHAATDTPPSVVHVHVCSDQNEEEEEEMERMKRP 180  
 DB 121 EDATPVINDLLVAISKLGPGKGYELHAATDTPPSVVHVHVCSDQNEEEEEEMERMKRP 180  
 QY 181 KPKIOTRRPEYTPIHLS 198  
 DB 181 KPKIOTRRPEYTPIHLS 198  
 RESULT 4  
 ID AAE18914  
 XX AAE18914 standard; Protein; 197 AA.  
 AC AAE18914;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated protein #1.  
 XX  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.

XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 XX Williams SR, Rothermel B;  
 XX  
 DR WPI: 2002-179698/23.  
 DR N-PSDB; AAD30155.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 XX (MCIP) binding, expression or phosphorylation, useful for treating  
 XX cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 XX calcineurin and a test compound -  
 XX  
 PS Example 1; Page 157-158; 174pp; English.  
 XX  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 XX and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 XX complex with the catalytic subunit of calcineurin and increased levels  
 XX of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 XX transcription of certain target genes. The invention also relates to  
 XX methods for identifying modulators of MCIP binding, expression or  
 XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 XX may be used for treating cardiac hypertrophy and heart failure.  
 XX Antibodies to MCIP can be used in characterising the MCIP content of  
 XX healthy and diseased tissues and subsequently for determining the  
 XX presence or absence of cardiomyopathy or as predictor of heart disease.  
 XX The present sequence is human MCIP associated protein.  
 XX Note: This sequence has been described as murine splice variant in  
 XX the specification, however the sequence seems to be a MCIP associated  
 XX protein.  
 XX  
 SQ Sequence 197 AA;  
 Query Match 83.3%; Score 872.5; DB 23; Length 197;  
 Best Local Similarity 89.4%; Pred. No. 6.2e-89;  
 Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;  
 QY 11 SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSRVRINFNPLSADARL 70  
 DB 11 SLLIACVANSDFSESETRAKFESLFRITYDKDTTFQYFKSRVRINFNPLSADARL 70  
 QY 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130  
 DB 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130  
 QY 131 LLYAISKLGPGEKVELHAATDTPPSVVHVHVCSDQNEEEEEEMERMKRPKPIQTRRP 190  
 DB 131 LLYAISKLGPGEKVELHAATDTPPSVVHVHVCSDQNEEEEEEMERMKRPKPIQTRRP 190  
 QY 191 EYTPIHLS 198  
 DB 190 EYTPIHLS 197  
 RESULT 5  
 ID AAW73898  
 XX AAW73898 standard; Protein; 171 AA.  
 AC AAW73898;  
 XX  
 DT 09-APR-1999 (first entry)  
 XX  
 DE Human Down's Syndrome critical region 1 protein.  
 XX  
 KW DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;  
 KW Central Nervous System development; mental retardation; heart defect.



QY 181 KPXIOTRRPEYTP 194  
Db 217 KPXIOTRRPGLPP 230

RESULT 7  
AAM50760  
ID AAM50760 standard; Protein; 255 AA.  
XX AAM50760;  
XX  
XX 18-APR-2002 (first entry)  
XX  
XX Human Down syndrome critical region 1-like 1 protein.  
XX  
XX Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;  
KW Down syndrome; Alzheimer's disease; dementia; transgenic;  
KW neuroprotective; nontropic; anticonvulsant; diagnosis; therapy.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 11  
FT /note= "encoded by TNG"  
FT Domain 129..199  
FT /note= "nucleic acid-binding domain"  
FT Region 169..176  
FT /note= "ISPPXSP box"  
XX  
XX WO200204513-A2.  
XX  
XX 17-JAN-2002.  
XX  
XX 11-JUL-2001; 2001WO-US21982.  
XX  
XX 11-JUL-2000; 2000US-0614474.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Loring JF, Tingley DW, Edwards CM, Streeter DG;  
XX WPI; 2002-164633/21.  
XX N-PSDB; ABA91457.  
XX  
XX Novel Down syndrome critical region 1-like 1 protein and nucleic acid  
PT encoding the protein useful for diagnosis and treatment of Alzheimer's  
PT disease, Down syndrome and other forms of dementia  
XX  
XX Claim 1; Page 47-49; 54pp; English.  
XX  
XX The present sequence is that of novel Down syndrome critical region  
CC 1-like 1 protein (DSCR1L1 alpha). Northern analysis indicated  
CC expression of DSCR1L1 alpha in various libraries, with the highest  
CC abundance in tissues from the nervous system, including tissues  
CC associated with schizophrenia, Huntington's disease, epilepsy and  
CC amyotrophic lateral sclerosis. An absence of DSCR1L1 expression  
CC was observed in 7 of 8 libraries from subjects with Alzheimer's  
CC disease. A claimed method for detecting differential expression of  
CC a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down  
CC syndrome, Alzheimer's disease and other forms of dementia. The  
CC purified DSCR1L1 alpha protein, or a portion of it, can be used  
CC to identify ligands which specifically bind the protein, such as  
CC a DNA or RNA molecule, peptide nucleic acid, peptide, protein,  
CC mimetic, agonist, antagonist, antibody, immunoglobulin, inhibitor  
CC or drug, for use as a therapeutic. The protein can also be used to  
CC raise antibodies useful in diagnosis and therapy.  
XX  
XX Sequence 255 AA;  
SQ

Query Match 59.2%; Score 620.5; DB 23; Length 255;  
Best Local Similarity 62.4%; Pred. No. 1.2e-60;  
Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

QY 5 DLODPSATIACHLDPRVFDGLCKRAKESLFRYDQDTTFQYFKSKRVIRINFSNPLSA 64  
Db 63 DFNDLNSLPACNVHOSVFEGEESKEKPEGLFRFYDDCVTFQLFKSFRVRINFSNPKSA 122  
QY 65 ADARLRHKTEFLGKEMKLYFAQTLHIGSS- ---HLAPPNPKQFLISPPASPPVGMKOV 120  
Db 123 ARARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVGMQPI 182  
QY 121 EDATPVINYDLYAISKLGEGEYELHAATDPTSPVVHVCSQDNEEEEMERMKRP 180  
Db 183 NDATPVINYDLYAVAKLGEGEYELHAGTSTSPVVHVCDSDIEEDPK- ---TSP 237  
QY 181 KPXIOTRRPEYTP 194  
Db 238 KPXIOTRRPGLPP 251

RESULT 8  
AAE18913  
ID AAE18913 standard; Protein; 197 AA.  
XX AAE18913;  
XX  
XX 17-MAY-2002 (first entry)  
XX  
XX Mouse MCIP associated proein #4.  
XX  
XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; mouse.  
XX  
XX Mus musculus.  
XX  
XX WO200204491-A2.  
XX  
XX 17-JAN-2002.  
XX  
XX 06-JUL-2001; 2001WO-US21662.  
XX  
XX 07-JUL-2000; 2000US-216601P.  
PR 13-FEB-2001; 2001US-0782953.  
XX  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PA (WILLI) WILLIAMS S R.  
PA (ROTH/) ROTHERMEL B.  
XX  
XX Williams SR, Rothermel B;  
PI  
XX WPI; 2002-179698/23.  
XX N-PSDB; AAD30154.  
XX  
XX Screening for modulators of muscle calcineurin interacting protein  
PT (MCIP) binding, expression or phosphorylation, useful for treating  
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
PT calcineurin and a test compound -  
XX  
XX Disclosure; Page 154; 174pp; English.  
XX  
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
CC complex with the catalytic subunit of calcineurin and increased levels  
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
CC transcription of certain target genes. The invention also relates to  
CC methods for identifying modulators of MCIP binding, expression or  
CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or  
CC may be used for treating cardiac hypertrophy and heart failure.  
CC Antibodies to MCIP can be used in characterizing the MCIP content of  
CC healthy and diseased tissues and subsequently for determining the  
CC presence or absence of cardiomyopathy or as predictor of heart disease.  
CC The present sequence is mouse MCIP associated protein.  
CC Note: This sequence has been described as mouse MCIP2 encoding DNA in  
CC the specification, however the sequence seems to be a MCIP associated  
CC protein.  
XX



transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterizing the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is human MCIP associated protein. Note: This sequence has been described as splice variant of MCIP1 initiated by exon 4 in the specification, however the sequence seems to be a MCIP associated protein.

Sequence 241 AA;  
Query Match 56.9%; Score 596; DB 23; Length 241;  
Best Local Similarity 61.9%; Pred. No. 6.1e-58;  
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 EVDLQDLPSATIACHLDPVFDGLCRKAFESLFRDYDKDTTQYKSKFRVRINFSNP 61  
DB 38 EMDLSDLPTSLFACSVHEAVFEAREQERFEALTYYDDQVTFQKSFRRVRINFSKP 97  
QY 62 LSAADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHLAPPNDKQFLISPPASPVGW 117  
DB 98 EAAARARIELHETDFNGOKLXYFAQVQMSGEVRDKSYLLPPQVQKQFLISPPASPVGW 157  
QY 118 KOVEDATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVCSQDNEEEEMERM 177  
DB 158 KQSDAMPVINYDLYLCAVSKLGPEKVELHAGTSTPSVVHVCSQDNEEEET 211  
QY 178 KRPKPKIIQTRPE 191  
DB 212 KNPQKIAQTRRPD 225

RESULT 11  
AAE18918  
ID AAE18918 standard; Protein; 212 AA.  
XX AAE18918;  
XX 17-MAY-2002 (first entry)  
XX Human MCIP associated protein #5.  
XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
XX heart failure; cardiomyopathy; heart disease; human.  
XX Homo sapiens.  
XX WO200204491-A2.  
XX 17-JAN-2002.  
XX 06-JUL-2001; 2001WO-US21662.  
XX 07-JUL-2000; 2000US-216601P.  
XX 13-FEB-2001; 2001US-0782953.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX (WILL/) WILLIAMS S R.  
XX (ROTH/) ROTHERMEL B.  
XX Williams SR, Rothermel B;  
XX WPI; 2002-179698/23.  
XX N-PSDB; AAD30159.  
XX Screening for modulators of muscle calcineurin interacting protein  
XX (MCIP) binding, expression or phosphorylation, useful for treating  
XX cardiac hypertrophy or heart failure, comprises mixing MCIP,  
XX calcineurin and a test compound "

PS Disclosure; Page 171-172; 174pp; English.  
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical  
XX complex with the catalytic subunit of calcineurin and increased levels  
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate  
XX transcription of certain target genes. The invention also relates to  
XX methods for identifying modulators of MCIP binding, expression or  
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
XX may be used for treating cardiac hypertrophy and heart failure.  
XX Antibodies to MCIP can be used in characterizing the MCIP content of  
XX healthy and diseased tissues and subsequently for determining the  
XX presence or absence of cardiomyopathy or as predictor of heart disease.  
XX The present sequence is human MCIP associated protein described in the  
XX invention.  
XX Sequence 212 AA;  
SQ Query Match 56.8%; Score 595; DB 23; Length 212;  
Best Local Similarity 62.6%; Pred. No. 6.5e-58;  
Matches 119; Conservative 20; Mismatches 39; Indels 12; Gaps 2;

QY 2 EVDLQDLPSATIACHLDPVFDGLCRKAFESLFRDYDKDTTQYKSKFRVRINFSNP 61  
DB 19 EMDLSDLPTSLFACSVHEAVFEAREQERFEALTYYDDQVTFQKSFRRVRINFSKP 78  
QY 62 LSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPVGWQKVE 121  
DB 79 EAAARARIELHETDFNGOKLXYFAQ-----SYLLPPQVQKQFLISPPASPVGWQKSE 132  
QY 122 DATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVCSQDNEEEEMERMKRPK 181  
DB 133 DAMPVINYDLYLCAVSKLGPEKVELHAGTSTPSVVHVCSQDNEEEET-----TKRPK 186  
QY 182 PKIIQTRPE 191  
DB 187 OKIAQTRRPD 196

RESULT 12  
AAB58452  
ID AAB58452 standard; Protein; 142 AA.  
XX AAB58452;  
XX 14-MAR-2001 (first entry)  
XX Lung cancer associated polypeptide sequence SEQ ID 790.  
XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;  
XX cardioactive; immunomodulatory; muscular active; vulnary;  
XX gastrointestinal; nephrotropic; antiinfective; gynecological;  
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
XX proliferative disorder; wound healing; infectious disease.  
XX Homo sapiens.  
XX WO200055180-A2.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05918.  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
XX Ruben SM;  
XX WPI; 2000-587514/55.  
XX N-PSDB; AAF18328.  
XX

PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer -

Claim 11; Page 1325-1326; 1425pp; English.

XX polynucleotide sequences AA17982 - AA18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytoskeletal; cardiovascular;  
CC immunomodulatory; muscular active general; vulnerability; gastrointestinal  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AA18425 - AA18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterization of the polynucleotide and protein  
CC sequences.

SO Sequence 142 AA;

Query Match 47.4%; Score 497; DB 21; Length 142;  
Best Local Similarity 96.9%; Pred. No. 3.2e-47;  
Matches 95; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MERVVDIODLPATITACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSN 60  
DB 38 MERVVDIODLPATITACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSN 97  
OY 61 PLSSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAP 98  
DB 98 PLSSADARLRLHKTETFLGKEMKLYFAQTLLHIGSSHLAP 135

RESULT 13

AA001768  
ID AAG01768 standard; Protein; 111 AA.

XX AAG01768;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5849.

KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.

XX Homo sapiens.

OS EPI033401-A2.

PN EPI033401-A2.

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GBST ) GENSET.

PI Dumas Milne Edwards J, Duclerc A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC01774.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNA or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

SO Sequence 111 AA;

Query Match 41.9%; Score 439; DB 21; Length 111;  
Best Local Similarity 84.2%; Pred. No. 6.6e-41;  
Matches 85; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 11 SATACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSSADARLR 70  
DB 11 SLIACVANSDFSESETRAFESLFRITYDKDTTFQYFKSPKRVINFSNPLSSADARLR 70  
OY 71 LHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNPDKQFLISPPA 111  
DB 71 LHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNPDKQFLISPPA 111

RESULT 14

AB071467  
ID ABB71467 standard; Protein; 292 AA.

XX ABB71467;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 41193.

KM Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL15570.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX disclosure; SEQ ID NO 41193; 21pp + Sequence listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of



```
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA
CC sequences (ABU101840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ . Sequence      292 AA;

    Query Match          36.2%; Score 379.5; DB 22; Length 292;
    Best Local Similarity 40.7%; Pred. No. 1.2e-33;
    Matches 79; Conservative 42; Mismatches 60; Indels 13; Gaps

Oy   3 EVD---LQDLP SATIACHLDPRVFDVGLCRAKFESLFRITYDKDTTFQVFKSFKRVRNES 59
Db   105 EVDADSPDLLPTSIIVNIHSEVFANPELKHAMBELFTFSASATFWLRSFRLRVNYD 164

Oy   60 NPLSAADARLEHLKTEFLGKE-MKLYPFAQLT-HIGSGHLAPPNDPKQLISPPASPVGW 117
Db   165 NAIAAANAARIKHOFVEFNKKTVITCYPAQPVPVSNKNLQPPAPVKQFLISPPASPAGW 224

Oy   118 KOVEDATPVPINVYLLVAISKLGCEKVELHAATDTPSVVVHVHCSDQNEEEEEEEERM 177
Db   225 EPREGEPVLNHDLLALASLTPEGSELHPQSEDQPAIIVHTAML-----AETGPGLQV 279

Oy   178 KRPRPKIIOTRRPE 191
Db   280 KAP---IVQT KCPE 290

RESULT 15
ID ABB29495
AC ABB29495 standard; Peptide; 58 AA.
XX AC ABB29495;
XX XX
DT 01-FEB-2002 (first entry)
XX XX
DE Peptide #2146 encoded by breast cell single exon nucleic acid probe.
XX Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer.
XX XX
OS Homo sapiens.
XX OS
WO 2000157271-A2.
XX PN
XX PD
XX PF 30-JAN-2001; 2001WO-US00662.
XX PR
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX XX
XX PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX XX
PS Claim 27; SEQ ID NO 12463; 327pp + sequence listing; English.
XX PS
XX XX
XX CC The invention relates to a spatially-addressable set of single exon
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.5835 Seconds  
(without alignments)  
1651.161 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCDVSTLVACVVDV.....SPKPKIOTRRPGLPPSVN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL 21.\*  
2: sp\_archaea.\*  
3: sp\_bacteria.\*  
4: sp\_fungi.\*  
5: sp\_human.\*  
6: sp\_invertebrate.\*  
7: sp\_mammal.\*  
8: sp\_mhc.\*  
9: sp\_organelle.\*  
10: sp\_phage.\*  
11: sp\_plant.\*  
12: sp\_rodent.\*  
13: sp\_virus.\*  
14: sp\_vertebrate.\*  
15: sp\_unclassified.\*  
16: sp\_rvirus.\*  
17: sp\_bacteriapi.\*  
17: sp\_archaeap.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000	96.1	192	11 Q8VIP5	Q8vip5 mus musculus
2	926	89.0	243	11 Q8VIP4	Q8vip4 mus musculus
3	634.5	61.0	198	11 Q91WQ4	Q91wq4 mus musculus
4	588	56.5	171	4 Q96R03	Q96r03 homo sapien
5	133.5	12.8	315	3 Q9P5S9	Q9p5s9 neurospora
6	129	12.4	249	3 Q9F4A1	Q9f4a1 cryptococcus
7	95	9.1	455	11 Q920L8	Q920l8 mus musculus
8	95	9.1	535	11 Q9EQT7	Q9eqt7 mus musculus
9	93	8.9	1018	6 Q28106	Q28106 bos taurus
10	92	8.8	277	3 Q13703	Q13703 schizosach
11	91	8.7	704	5 Q17478	Q17478 hyalophora
12	90.5	8.7	1213	4 Q92923	Q92923 homo sapien
13	89	8.5	746	4 Q9BU60	Q9bu60 homo sapien
14	89	8.5	874	5 Q9VQL0	Q9vql0 drosophila
15	89	8.5	1021	4 Q15451	Q15451 homo sapien
16	89	8.5	1052	4 Q96FT1	Q96ft1 homo sapien

17	89	8.5	1251	4 Q15450	Q15450 homo sapien
18	88.5	8.5	465	4 Q96GY4	Q96gy4 homo sapien
19	88.5	8.5	501	5 Q9BKQ7	Q9bkq7 caenorhabdi
20	88.5	8.5	1130	4 Q96E12	Q96e12 homo sapien
21	88.5	8.5	1214	4 Q8TAQ2	Q8taq2 homo sapien
22	87.5	8.4	13055	5 Q09165	Q09165 caenorhabdi
23	86	8.3	1006	11 Q62901	Q62901 rattus norv
24	86	8.3	1066	4 Q9H2N5	Q9h2n5 homo sapien
25	85.5	8.2	320	5 Q95SK4	Q95sk4 drosophila
26	85.5	8.2	320	5 Q9VOR5	Q9vqr5 drosophila
27	85.5	8.2	811	2 Q59316	Q59316 clostridium
28	85.5	8.2	7576	2 Q9ZGA4	Q9zga4 streptomyc
29	85	8.2	878	5 Q9VZF8	Q9vzf8 drosophila
30	85	8.2	4025	4 Q9NR13	Q9nr13 homo sapien
31	84.5	8.1	212	10 Q9LYU3	Q9lyu3 arabidopsis
32	84.5	8.1	314	10 Q9LIV7	Q9liv7 oryza sativ
33	84.5	8.1	1114	5 Q9V180	Q9v180 drosophila
34	84.5	8.1	1222	16 Q98PR9	Q98pr9 mycoplasma
35	84.5	8.1	2701	4 Q9Y520	Q9y520 homo sapien
36	84.5	8.1	3938	11 Q88778	Q88778 rattus norv
37	84.5	8.1	5002	5 Q9VXR3	Q9vvr3 drosophila
38	84	8.1	246	10 Q23143	Q23143 arabidopsis
39	84	8.1	296	10 Q8W0W9	Q8w0w9 oryza sativ
40	84	8.1	671	10 Q9FXA2	Q9fxa2 arabidopsis
41	84	8.1	1557	5 Q96652	Q96652 drosophila
42	83.5	8.0	220	10 Q9FGA3	Q9fga3 arabidopsis
43	83.5	8.0	564	5 Q9NA93	Q9na93 caenorhabdi
44	83	8.0	160	17 Q9YD46	Q9y46 aeropyrum p
45	83	8.0	643	5 Q9XUT0	Q9xut0 caenorhabdi

## ALIGNMENTS

### RESULT 1

Q8VIP5 PRELIMINARY; PRT; 192 AA.  
AC Q8VIP5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Calcineurin inhibitory protein ZAKI-4.  
GN ZAKI-4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kanou Y., Miyazaki T., Seo H., Murata Y.;  
RT "calcineurin inhibitory protein ZAKI-4";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB061524; BAB71955.1; --  
SQ SEQUENCE 192 AA; 21540 MW; B2DB9819F4B36AE8 CRC64;

Query Match 96.1%; Score 1000; DB 11; Length 192;

Best Local Similarity 98.4%; Pred. No. 1e-90; 2; Indels 0; Gaps 0;

Matches 189; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 MDCDVSTLVACVVDVEVTNQEVKEKFEGLFRTYDECVTQFQKSFRRVRINFSPKSA 65

Db 1 MDCDVSTLVACVVDVEVTNQEVKEKFEGLFRTYDECVTQFQKSFRRVRINFSPKSA 60

Qy 66 RARIELHETQFRGKKLYFAQVOTPEETDQKHLAPPQAKQFLISPPSPPCWKQDIS 125

Db 61 RARIELHETQFRGKKLYFAQVOTPEETDQKHLAPPQAKQFLISPPSPPCWKQDIS 120

Qy 126 DATPVLNLDLLYAVAKLPGEGKEKELHAGTSTPSVWVHVCDSDMEEDDPKTSKPKEIIQ 185

Db 121 DATPVLKDYDLLYAVAKLPGEGKEKELHAGTSTPSVWVHVCDSDMEEDDPKTSKPKEIIQ 180

Qy 186 TRRPGLPPSVSN 197

Db 181 TRPGLPSPVSN 192

# RESULT 2

Q8VIP4 PRELIMINARY; PRT; 243 AA.

AC Q8VIP4; 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4 beta.  
 GN ZAKI-4 BETA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxId=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,  
 RT "Calcineurin inhibitory protein ZAKI-4."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB061525; BAB71956.1; -  
 DR InterPro; IPR000504; RNM\_rec\_moc.  
 DR Pfam; PF00076; rtm; 1.  
 SQ SEQUENCE 243 AA; 27332 MW; 2CD8C1A810291851 CRC64;

Query Match 89.0%; Score 926; DB 11; Length 243;  
 Best Local Similarity 94.1%; Pred. No. 2.6e-83;  
 Matches 176; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 11 STLAVCVDEVFPTNOEVEKEFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 70  
 Db 57 NSLFACNVHQSVEEESKEKFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 130  
 Db 117 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 190  
 Db 177 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 236

Qy 191 LPPSVSN 197  
 Db 237 LPPSVSN 243

RESULT 3  
 Q91WQ4 PRELIMINARY; PRT; 198 AA.

AC Q91WQ4; 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Down syndrome critical region homolog 1 (human).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxId=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013551; AAH13551.1; -  
 SQ SEQUENCE 198 AA; 22851 MW; F017C68F19ACC187 CRC64;

Query Match 61.0%; Score 634.5; DB 11; Length 198;  
 Best Local Similarity 64.4%; Pred. No. 1.1e-54;  
 Matches 121; Conservative 25; Mismatches 33; Indels 9; Gaps 2;

Qy 11 STLAVCVDEVFPTNOEVEKEFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 70  
 Db 57 NSLFACNVHQSVEEESKEKFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 130  
 Db 117 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 190  
 Db 177 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 236

Qy 191 LPPSVSN 197  
 Db 237 LPPSVSN 243

Db 11 SSLIACVANDVSESETRAKFESLFRITYDQDTTFQYFKSFRRVRINFSNPLSAAADARLR 70

Qy 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 130

Db 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 126

Qy 131 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 190  
 Db 127 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 186

Qy 186 TRPGLPSPVSN 193  
 Db 187 TRPGLPSPVSN 194

RESULT 4  
 Q96R03 PRELIMINARY; PRT; 171 AA.

AC Q96R03; 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Down syndrome critical region protein 1.  
 GN DSCR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxId=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,  
 RA Qiang B.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF400429; AAK92478.1; -  
 SQ SEQUENCE 171 AA; 19823 MW; 133001A8AEDF0BD9 CRC64;

Query Match 56.5%; Score 588; DB 4; Length 171;  
 Best Local Similarity 66.5%; Pred. No. 3.6e-50;  
 Matches 113; Conservative 22; Mismatches 27; Indels 8; Gaps 2;

Qy 28 VKEKEGELFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIEHETQFRGKKLKIYFAO 87  
 Db 2 VYAKESLFRITYDQDTTFQYFKSFRRVRINFSNPLSAAADARLRHETQFRGKKLKIYFAO 61

Qy 88 VOTPETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPVANVLLVAVAKLGEGEK 147  
 Db 62 TLHIGSS---HLAPPNDKQFLISPPSPVGMKPISDATPVINVLLVAVAKLGEGEK 117

Qy 148 YELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 193  
 Db 118 YELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 167

RESULT 5  
 Q9P5S9 PRELIMINARY; PRT; 315 AA.

AC Q9P5S9; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE Nebula related protein.  
 GN B5022.230.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OC NCBI\_TaxId=5141;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hohenberg J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF400429; AAK92478.1; -  
 SQ SEQUENCE FROM N.A.

Qy 11 STLAVCVDEVFPTNOEVEKEFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 70  
 Db 57 NSLFACNVHQSVEEESKEKFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 130  
 Db 117 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 190  
 Db 177 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 236

Qy 191 LPPSVSN 197  
 Db 237 LPPSVSN 243

Qy 11 STLAVCVDEVFPTNOEVEKEFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 70  
 Db 57 NSLFACNVHQSVEEESKEKFEGLFRITYDECYTFQLFKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 130  
 Db 117 LHETQFRGKKLKIYFAOVQPTETDGDKLHAPPOAKOFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 190  
 Db 177 LNVDLVAVAKLGEGEYELHAGTSTPSVVAVHCDSDMEEDPKTSPPKIIQTRRPG 236

Qy 191 LPPSVSN 197  
 Db 237 LPPSVSN 243

```

RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355932; CAB91442.1; -.
SQ SEQUENCE 315 AA; 34194 MW; 76792987653JAB7 CRC64;

  Query Match      12.8%; Score 133.5; DB 3; Length 315;
  Best Local Similarity 37.1%; Pred. No. 4.9e-05;
  Matches 36; Conservative 12; Mismatches 46; Indels 3; Gaps 2;

QY 49 KSPRRVRINFSHPKSAARARIELHETQFRGKGLKLYFAQVQTPETDGDKLHLAPPOPAKQ 108
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 101 KSPRRRIIVTFDQAAIAVRSVWDGEAIIIGRCRVRVFGOPTIDVSAADKHLALPDAGKL 160
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 109 FLSPSPSPVGVKHP-ISA--TPVLNLYLLYAVAKL 142
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 FLSPSPSPVGHQWEDAPMTVMHAEADLAELAKL 197
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
Q9P4A1 PRELIMINARY; PRT; 249 AA.
AC Q9P4A1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Calcineurin-binding protein.
GN CBP1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=H99;
RX Goriach J., Fox D.S., Cutler N.S., Cox G.M., Perfect J.R., Heitman J.;
RA "Identification and characterization of a highly conserved calcineurin
RT binding protein, CBP/calciopressin, in Cryptococcus neoformans.";
RL EMBO J. 19:3618-3629(2000).
DR EMBL; AF230799; AAF63734.1; -.
SQ SEQUENCE 249 AA; 27293 MW; 7B757183FA5EBF09 CRC64;

  Query Match      12.4%; Score 129; DB 3; Length 249;
  Best Local Similarity 34.0%; Pred. No. 0.0001;
  Matches 35; Conservative 15; Mismatches 35; Indels 18; Gaps 3;

QY 81 LKLYFAQVQTPETDGDKLHLAPPOPAKQFLISPPSPVGVKWPISDATP---VLNVDLLY 137
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 LRLHLYLPTTLNPDPTATTHLAPPPLPHNFIISPPGSPGWEPAEAPNRILPEDLQR 166

QY 138 AVAKLPGGKGYELHAGTESTP-----SVVVHVCDSDMBE 171
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 AL-----ETLEINSGSKADGDKKEIILDEGVRVQVEDTTKQE 203

RESULT 7
Q920L8 PRELIMINARY; PRT; 455 AA.
AC Q920L8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fca/m receptor (fragment).
GN FCAFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_SEQUENCE FROM N.A.
RP Shimizu Y., Honda S., Yotsumoto K., Tahara-Hanaoka S., Eyre H.J.,
RA Sutherland G.D., Endo Y., Shibuya K., Koyama A., Nakauchi H.,
RA Shibuya A.;

```

```

RT "Fca/m receptor is a single gene-family member closely related to
RL polymeric immunoglobulin receptor on chromosome 1.";
DR EMBL; AB071978; BAB71750.1; -.
SQ SEQUENCE 315 AA; 34194 MW; 76792987653JAB7 CRC64;

  Query Match      9.1%; Score 95; DB 11; Length 455;
  Best Local Similarity 25.0%; Pred. No. 0.48;
  Matches 29; Conservative 19; Mismatches 50; Indels 18; Gaps 4;

QY 87 QVQTPETDGDK-----LHLAPPOPAKQFLISPPS---SPVGVKWPISDATPVLNVDLLYAV 139
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 QGTTPTDGPREDTDRVSPAPRKTGTTRPSALISEHVTWETLQDKTEVSKQQLHSL 302

QY 140 AKLGPGEKYE-----LHAGTESPTSVVVHVCDSDMBEEDPKTSPKPKIOTRRP 189
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 BELSPAPSAQTLNATCLEVASEGRSI-----DGSLENTTESSPPTPSQLSVAGP 353

RESULT 8
Q9EQ7 PRELIMINARY; PRT; 535 AA.
AC Q9EQ7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fca/m receptor.
GN FCAFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_SEQUENCE FROM N.A.
RP MEDLINE=21170225; PubMed=11062505;
RX Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
RA Hironaka T., Eyre H.J., Sutherland G.R., Endo Y., Fujita T.,
RA Miyabayashi T., Sakano S., Tsuji T., Nakayama E., Phillips J.H.,
RA Lanier L.L., Nakauchi H.;
RT "Fca/m receptor mediates endocytosis of Igm-coated microbe.";
RL Nat. Immunol. 1:441-446(2000).
DR EMBL; AB048834; BAB17312.1; -.
DR MGD; MGI:1927803; Fcamr.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000508; SigPase.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00410; Ig_Like; 1.
DR PROSITE; PS00501; SPASE_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B7C70151C75 CRC64;

  Query Match      9.1%; Score 95; DB 11; Length 535;
  Best Local Similarity 25.0%; Pred. No. 0.59;
  Matches 29; Conservative 19; Mismatches 50; Indels 18; Gaps 4;

QY 87 QVQTPETDGDK-----LHLAPPOPAKQFLISPPS---SPVGVKWPISDATPVLNVDLLYAV 139
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 QGTTPTDGPREDTDRVSPAPRKTGTTRPSALISEHVTWETLQDKTEVSKQQLHSL 382

QY 140 AKLGPGEKYE-----LHAGTESPTSVVVHVCDSDMBEEDPKTSPKPKIOTRRP 189
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 383 BELSPAPSAQTLNATCLEVASEGRSI-----DGSLENTTESSPPTPSQLSVAGP 433

```

## RESULT 9

028106 PRELIMINARY; PRT; 1018 AA.

AC 028106; 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE F3/F11/contactin precursor.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxId=9913;  
 RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;  
 RX MEDLINE=95369697; PubMed=7642103;  
 RA Watanabe K., Shimazaki K., Hosoya H., Fukumachi F., Takenawa T.;  
 RT "Cloning of the DNA encoding neural adhesion molecule F3 from bovine  
 brain";  
 RL Gene 160:245-248 (1995).

DR EMBL; D32135; BAA06861.1; -  
 DR HSSP; P08921; 1A7B.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR003598; IG\_C2.  
 DR InterPro; IPR003600; IG\_1like.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR002052; N6\_Mtase.  
 DR Pfam; PF00041; Ig; 4.  
 DR Pfam; PF00047; Ig; 6.  
 DR SMART; SMO0060; FN3; 4.  
 DR SMART; SMO0408; IGc2; 4.  
 DR SMART; SMO0410; IG\_1like; 2.  
 DR PROSITE; PS00092; N6\_MTASE; UNKNOWN\_1.  
 DR Immunoglobulin domain; signal.  
 FT SIGNAL 1 20  
 KM SEQUENCE 1018 AA; 113384 MW; 551FC10ED7920341 CRC64;

Query Match 8.9%; Score 93; DB 6; Length 1018;

Best Local Similarity 21.4%; Pred. No. 2.1; Matches 37; Conservative 30; Mismatches 70; Indels 36; Gaps 7;

OY 9 DVSTLVACVDEVEFTNOEYKEKEFEGLFRTYDECVTFOLEKSFRRVYRINFSHPKSAAR 68  
 DB 806 DAPSEPTAVGVVLSSEISVMEH-----VVEKIVESY-QIRYASHKKAAR 856  
 OY 69 IELHETQFRGKKLKL-----YFAQVOTPETDGDKHLAPPOPAKQFLT--SPSSFP--- 118  
 DB 857 VOVASQGEYSARLEMLLPDIOYFVEVRACNSAG---CGPPSDMTETFTKKAPSPRPRII 912  
 OY 119 -----VGK---PISDAPVINYDLVAVAKLGSEKELHAGTESTP 158  
 DB 913 SSVRSGRYIITWDHVVALSNSTVGVKLVKRPDGDHDKLSTYTKHSIEVP 965

## RESULT 10

013703 PRELIMINARY; PRT; 277 AA.

AC 013703; 01-JAN-1999 (TREMblrel. 09, Created)  
 DT 01-JAN-1999 (TREMblrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TREMblrel. 09, Last annotation update)  
 DE Hypothetical 29.6 kDa protein C135.04C in chromosome I.  
 GN SFAC135.04C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC NCBI\_TaxId=4996;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;  
 RA Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;  
 RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z99091; CAB1767.1; -  
 KW Hypothetical protein.  
 FT DOMAIN 267 270  
 SQ SEQUENCE 277 AA; 29622 MW; 540CF08359A06546 CRC64;

Query Match 8.8%; Score 92; DB 3; Length 277;  
 Best Local Similarity 22.0%; Pred. No. 0.52;  
 Matches 45; Conservative 24; Mismatches 76; Indels 60; Gaps 7;

OY 42 CYTFOLEFSFRVRINF-----SHPSAR-----ARILHETQFRGK 79  
 DB 2 CFFRLVSSFAILRLSIFLSIKIFCKRLALQVANPQRKAKTPSNHATEBLQOSTNST 61  
 OY 80 KKLKLYFAOVOTPETDGDKHLAP--OPAKQFLSPSSPSPVGMKP-----ISDA 127  
 DB 62 TLPTGEAAVETNASHETSRALPTTSPASLSISPTKSAVSSBPVADVKSISTPA 121  
 OY 128 TPVLNYDLVAVAKLGSEKYE-----LHAGTESTPSV-----VHVCDSDME 171  
 DB 122 APQLN-----SPHSYETPTTPTSTITENLPTIDPTRSTRSSSHIQSLSPESKO 171  
 OY 172 EDPKTSPPKPIQTRRPGLPSPVS 196  
 DB 172 TSDGHRPPSPSTSTTSTSDPSVA 196

## RESULT 11

017478 PRELIMINARY; PRT; 704 AA.

AC 017478; 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Arylphorin precursor.  
 OS Arylphorin cecropia (Cecropia moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Bombycoidea; Saturniidae; Saturniinae; Atteciini; Hyalophora.  
 OC NCBI\_TaxId=7123;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FAT BODY;  
 RA Massey H.C. Jr.;  
 RT "The Evolution of Insect Storage Proteins: Driven by Composition and  
 RT Constrained by Sequence."  
 RT Thesis (1995), University of Pennsylvania.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FAT BODY;  
 RX MEDLINE=98329402; PubMed=9664700;  
 RA Bumester T., Massey H.C. Jr., Zakharin S.O., Benes H.;  
 RT "The evolution of hexamerins and the phylogeny of insects."  
 DR J. Mol. Evol. 47:93-108 (1998).  
 DR EMBL; AF032396; AAB86644.1; -  
 DR HSSP; P04253; 1ILA.  
 DR InterPro; IPR000896; Hemocyanin.  
 DR InterPro; IPR005203; hemocyanin\_C.  
 DR InterPro; IPR005204; hemocyanin\_N.  
 DR Pfam; PF00372; hemocyanin\_1.  
 DR Pfam; PF03723; hemocyanin\_C; 1.  
 DR Pfam; PF03722; hemocyanin\_N; 1.  
 DR PRINTS; PR00187; HAEMOCYANIN.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; 1.  
 KM SIGNAL 1 16  
 FT SIGNAL 17 704  
 FT CHAIN 17 704  
 SQ SEQUENCE 704 AA; 83788 MW; C1805EDA0ACD4626 CRC64;

Query Match 8.7%; Score 91; DB 5; Length 704;  
 Best Local Similarity 27.0%; Pred. No. 2.1;

```
Matches 40; Conservative 24; Mismatches 56; Indels 28; Gaps 6;
QY 7 DCVSTLVACVVDVVEFTNVEKFKGLPRT-----YDECVTFO-----LPKSFR 52
Db 64 DYDVE-----ANIDYTNKKAVEFLKYLTKYFEPFIFPHKLRBEAIALPHLY 117
QY 53 RVIRNFSPKSAARIELEHETOPRGKKLXYFAQVOTPETDCKLHLAPQPAKQFLIS 112
Db 118 YAKDFETFKSAAPARVHLNEGQP---LYAYYIAVIQNRDTHGVLDP-APYEVYVPPQFVN 173
QY 113 PPSPPVGVKPISDA-----TPVLNYDL 136
Db 174 MDITVRIYRTKMDGILHPTKAINYGIV 201

RESULT 12
Q92923
ID Q92923 PRELIMINARY; PRT; 1213 AA.
AC Q92923;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SWI/SNF complex 170 kDa subunit.
GN BAF170.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96397413; PubMed=8804307;
RA Wang W., Xue Y., Zhou S., Kuo A., Cairns B.R., Crabtree G.R.;
RT "Diversity and specialization of mammalian SWI/SNF complexes.";
RL Genes Dev. 10:2117-2130(1996).
[2]
RP SEQUENCE FROM N.A.
RA Wang W., Cote J., Xue Y., Zhou S., Khavari P.A., Biggar S.R.,
RA Muchardt C., Kaipana G.V., Goff S.P., Yaniv M., Workman J.L.,
RA Crabtree G.R.;
RT "Purification and biochemical heterogeneity of the mammalian SWI-SNF
RT complex.";
RL EMBO J. 15:0-0(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; U66616; AAC50694.1; -.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00298; CHROMO; 1.
DR SMART; SM00395; SANT; 1.
DR PROSITE; PS0090; MYB 3; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1213 AA; 132732 MW; D2139FD7E7CB5235 CRC64;

Query Match
Best Local Similarity 21.0%; Score 90.5; DB 4; Length 1213;
Matches 43; Conservative 26; Mismatches 69; Indels 67; Gaps 9;
QY 10 VSTLVACVVDVVEFTNVEKFKGLPRTYDECVTFOFKSFRVRVNFSPKSAARARI 69
Db 884 IKSLVALLVETOM-----KKLEIKLRHFELETTM-----DREREAL 920
QY 70 ELHETOPRGKKLXYFAQVOTPETDCKLHL-----APPOPAKQFLISPPSPVGVK 122
Db 921 EYQEQQLADRQAFHNEQLKYPEWRARQHQFQMHQOQQOQPPPA-----LPPGSQPI--P 973
QY 123 PISDATPVLNYDLLYAVA-----KLGRCEKVELHAGTESTPVSVVHVCDSD 168
Db 974 PTGAAGPAPVHGLAVAPASVVPAPAGSGAPPGSLGRSEIQIG-QAGSTRGP----- 1022
QY 169 MESEEDPKTSPKPKIIQTRRGLPP 193
```

```
Db 1023 --QQQQPAGAPQPGAVP---PGVPP 1042

RESULT 13
Q9BU60
ID Q9BU60 PRELIMINARY; PRT; 746 AA.
AC Q9BU60;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 79.2 kDa protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC002875; AA02875.1; -.
DR InterPro; IPR002965; P-rich_extensn.
DR PRINTS; PR01217; PRICHEXTENSN.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 746 AA; 79187 MW; 35A186C5A7ADEA8B CRC64;

Query Match
Best Local Similarity 8.5%; Score 89; DB 4; Length 746;
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;
QY 101 APPQPAKQELISPPSPVGVKPISDATPVLNYDLVAVAKLPGKVELHAGTESTPSV 160
Db 592 APPPTLPALPPSPSP-KVQPEPEP-----GLLEVEE--PGTEEE--RGAUTAPT 642
QY 161 VHVCDSDME---EBEDPKTSPKPKIIQTRRGLPPSV 195
Db 643 APEALPSQGEVEREGESPAAGPPQQLVEEPSAPPTL 680

RESULT 14
Q9VQLO
ID Q9VQLO PRELIMINARY; PRT; 874 AA.
AC Q9VQLO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CGI7258 protein.
GN CGI7258.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champé M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fieichmann W.,  
 RA Foster C., Gabrielsen A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Krat C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Lavieky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,  
 RA Palazzolo M., Pictman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003581; AAF51156.1; -  
 DR F1Base: FBgn0031496; CG17258.  
 SQ SEQUENCE 874 AA; 103694 MW; 5F56D1C0E7A01D9A CRC64;

Query Match 8.5%; Score 89; DB 5; Length 874;  
 Best Local Similarity 21.0%; Pred. No. 42;  
 Matches 50; Conservative 28; Mismatches 74; Indels 86; Gaps 9;

OY 6 MDCDVSTLVACVVDVEFTNOEYKEFEGI-FRTYDECYTFOLFKS-FRRVRI----- 56  
 DB 201 MDCWTEQGAQVVEIPAFIEBOV-ISFDALFVRNFTKQVQADSKXILKCRTPSPILTST 259  
 OY 57 -NFSHPKSAARARIELHETQFRGK-----LK----- 82  
 DB 260 EEFRRPMRAAAROMLHVTRVHEEDLLOENKQAOQPEGEPSLSKRESSTCSTSPGIDS 319  
 OY 83 -----LYAQVOTPPTDDGKHLAPPOPAK--QFLISPPSPVGMKPISDATPVL 131  
 DB 320 VTSDLAEALYELQLEAAKENEELDKELVPPKLVQFQLPKKSPTLTTHFVGDS---- 375  
 OY 132 NYDLVAVAKLGGEKELHAGTSTPSVYVHCSDMEEDDPKTPSPKXIQTTRP 189  
 DB 376 -----EAEPEQEPILIRI-----ELPKVYVAKPEBELKQF 406

RESULT 15

015451 PRELIMINARY: PRT; 1021 AA.

AC 015451;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Proline and glutamic acid rich nuclear protein isoform  
 DE (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Vadlamudi R.K., Chung J.K., Shin J.;  
 RT "p160.2 an isoform of proline and glutamic acid rich nuclear protein  
 RT p160.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U88154; AAC17709.1; -  
 DR InterPro: IPR002965; P rich extenan.  
 DR PRINTS: PRO1217; PRICHKXTNEN.  
 KW Nuclear protein.  
 FT NON\_TER 1 1

SQ SEQUENCE 1021 AA; 109136 MW; C8951B4D67698D3 CRC64;

Query Match 8.5%; Score 89; DB 4; Length 1021;  
 Best Local Similarity 29.6%; Pred. No. 5.1;  
 Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

OY 101 APPQAKFLISPPSPVGMKPISDATPVUNYDLVAVAKLGGEKELHAGTSTPSV 160  
 DB 867 APPPTLPALPPSPSP-KVQPEPEPEP---GLLLEVEE--PQTEBE--RGADTAFTL 917  
 OY 161 VHVCDSDME--EEDDPKTPSPKXIQTTRRGLPPSV 195  
 DB 918 APBALPSQGEYVERGEGSPAAQPPQELVVEEPPSAFTL 955

Search completed: December 11, 2002, 11:41:09  
 Job time : 27.5835 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3086 Seconds  
(without alignments)  
1117.976 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCDSTLVACVDV.....SPKPIQTRRPELPSPVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	1 CCP2_MOUSE	Q9jhg2 mus musculus
2	1007	96.7	197	1 CCP2_HUMAN	Q14206 homo sapien
3	663.5	63.7	239	1 CCP3_MOUSE	Q9jkk0 mus musculus
4	658.5	63.3	241	1 CCP3_HUMAN	Q9uka8 mus musculus
5	641	61.6	197	1 CCP1_CRIGR	O35847 cricetus
6	622	59.8	197	1 CCP1_HUMAN	P53805 homo sapien
7	615.5	59.1	198	1 CCP1_MOUSE	Q9jhg6 mus musculus
8	404	38.8	292	1 NLA_DROME	Q9xz18 drosophila
9	322.5	31.0	207	1 CCP1_CAEEL	P53806 caenorhabdi
10	107.5	10.3	163	1 YAPF_SCHPO	Q09791 schizosacch
11	101	9.7	454	1 YAPF_HUMAN	P46937 homo sapien
12	90	8.6	526	1 CAP_YEAST	P17555 saccharomyc
13	87.5	8.4	211	1 RCN1_YEAST	P36054 saccharomyc
14	84.5	8.1	1827	1 FM14_HUMAN	P11137 homo sapien
15	83	8.0	1206	1 FM14_MOUSE	Q05859 mus musculus
16	83	8.0	1232	1 Y005_CAEEL	P34643 caenorhabdi
17	83	8.0	1468	1 FMN1_MOUSE	Q09791 schizosacch
18	82.5	7.9	1377	1 NE01_RAT	P46937 homo sapien
19	82	7.9	2464	1 MAPB_MOUSE	Q05860 mus musculus
20	82	7.9	3703	1 ABF1_HUMAN	P97603 rattus norv
21	81.5	7.8	3726	1 ABF1_MOUSE	Q14873 mus musculus
22	81	7.8	559	1 ECN1_MOUSE	Q15911 homo sapien
23	81	7.8	2459	1 MAPB_RAT	Q61329 mus musculus
24	79	7.6	909	1 CNG4_HUMAN	Q61508 mus musculus
25	78.5	7.5	497	1 HME5_DROME	P15205 rattus norv
26	78.5	7.5	836	1 NOT3_YEAST	Q14028 homo sapien
27	78	7.5	1018	1 CONT_HUMAN	P18488 drosophila
28	78	7.5	2468	1 MAPB_HUMAN	P06102 saccharomyc
29	77.5	7.4	508	1 V56K_PLRV1	Q12860 homo sapien
30	77.5	7.4	928	1 CHS2_EXODE	P17525 potato leaf
31	77.5	7.4	1461	1 NE01_HUMAN	P30601 exophiala d
32	77	7.4	834	1 V2A_PSVJ	Q92859 homo sapien
33	77	7.4	1258	1 GLI2_HUMAN	P28727 peanut stun
					P10070 homo sapien

34 77 7.4 1302 1 NRG\_DROME P20241 drosophila  
35 76.5 7.3 397 1 A2BP\_HUMAN Q9nwb1 homo sapien  
36 76.5 7.3 452 1 ODO2\_SCHPO O94681 schizosacch  
37 76.5 7.3 1733 1 VNUA\_PRVKA P33485 pseudosacch  
38 76 7.3 551 1 CAP\_SCHPO P36621 schizosacch  
39 76 7.3 728 1 P85B\_HUMAN O00459 homo sapien  
40 76 7.3 793 1 S3A1\_HUMAN Q15459 homo sapien  
41 76 7.3 857 1 NFM\_CHICK P16053 gallus gall  
42 76 7.3 870 1 Y563\_HUMAN O60309 homo sapien  
43 76 7.3 930 1 SM6C\_HUMAN Q9h1t2 homo sapien  
44 76 7.3 1493 1 NE01\_MOUSE P97798 mus musculus  
45 75.5 7.3 614 1 ACES\_RAT P37136 rattus norv

#### ALIGNMENTS

RESULT 1  
CCP2\_MOUSE  
ID CCP2\_MOUSE STANDARD; PRT; 197 AA.  
AC Q9JHG2;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)  
DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).  
GN DSCR1L1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20187590; PubMed=10722714;  
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,  
RA Williams R.S.;  
RT "A protein encoded within the Down syndrome critical region is  
RT enriched in striated muscles and inhibits calcineurin signaling.";  
RL J. Biol. Chem. 275:8719-8725(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BA1B/c; TISSUE=Brain;  
RX MEDLINE=20534792; PubMed=11080588;  
RA Striappoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;  
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:  
RT conserved synteny with the human orthologous genes.";  
RL Gene 257:233-232(2000).  
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
CC by binding to the catalytic domain of calcineurin A. Could play a  
CC role during central nervous system development.  
CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle  
CC and brain. Lower expression in all other tissues.  
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AF237791; AAF63487.1; -;  
CC EMBL; AF237887; AAF62538.1; -;  
CC MGD; MGI:1858219; Dscr1l1.  
SQ SEQUENCE 197 AA; 22025 MW; CEI306B7B03E70F1 CRC64;  
Query Match 100.0%; Score 1041; DB 1; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1.le-83;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPAPSMDCDSTLVACVDVVEFTNQVKPEGLFRITYDECVTFQLPKSFRRVINFSSH 60  
|||||

```

Db      1  MPASMDCVSTLVACVDEVFPTNOEVEKEBGLFRITVDECVTQLFSPRRVAINFSH 60
Oy      61  PPSAARARIELEHETOPRGKKLKYFAOVOTPETDGDKLHAPPAKOPFLISPPSSPPVG 120
Db      61  PPSAARARIELEHETOPRGKKLKYFAOVOTPETDGDKLHAPPAKOPFLISPPSSPPVG 120
Oy      121  WKPISDATPVNLVLDLVAVALGPGKEKELHAGTESTPSVVVHVCDSDMEBEDPKTSPK 180
Db      121  WKPISDATPVNLVLDLVAVALGPGKEKELHAGTESTPSVVVHVCDSDMEBEDPKTSPK 180
Oy      181  PKIOTRRRGLPPSVSN 197
Db      181  PKIOTRRRGLPPSVSN 197

RESULT 2
CCP2_HUMAN  STANDARD;  PRT;  197 AA.
ID  CCP2_HUMAN
AC  Q14206;
DT  30-MAY-2000 (Rel. 39, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DE  15-UN-2002 (Rel. 41, Last annotation update)
DE  Calcipressin 2 (thyroid hormone-responsive protein ZAKI-4) (Down
DE  syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE  interacting protein 2) (MCIP2).
GN  DSCR11 OR ZAKI4.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Fibroblast;
RX  MEDLINE=96278928; PubMed=8662924;
RA  Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA  Yamazaki H., Seo H.;
RT  "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT  in human skin fibroblasts."
RJ  J. Biol. Chem. 271:14567-14571(1996).
CC  -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC  by binding to the catalytic domain of calcineurin A. Could play a
CC  role during central nervous system development.
CC  -1- TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC  and skeletal muscle but not in placenta, lung, kidney and
CC  pancreas.
CC  -1- INDUCTION: By thyroid hormone.
CC  -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.emb-eb.ch/announce/
CC  or send an email to license@eb-sib.ch).
CC  EMBL: D83407; BAA11911.1; ALT_INIT.
DR  Genew: HGNC:3041; DSCR11.
DR  MIM: 604876;
SQ  SEQUENCE 197 AA; 21955 MW; AF39735F6661C8E0 CRC64;

Query Match 96.7%; Score 1007; DB 1; Length 197;
Best Local Similarity 96.4%; Pred. No. 1e-80;
Matches 190; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      121  WKPISDATPVNLVLDLVAVALGPGKEKELHAGTESTPSVVVHVCDSDMEBEDPKTSPK 180
Db      121  WOPINDATPVNLVLDLVAVALGPGKEKELHAGTESTPSVVVHVCDSDMEBEDPKTSPK 180
Oy      181  PKIOTRRRGLPPSVSN 197
Db      181  PKIOTRRRGLPPSVSN 197

RESULT 3
CCP3_MOUSE  STANDARD;  PRT;  239 AA.
ID  CCP3_MOUSE
AC  Q9UKO0; Q9CXB7;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE  (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN  DSCR12.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BA1B/c; TISSUE=Brain;
RA  Strippoli P., Petroni M., Lenzi L., Carinci P., Zannotti M.;
RT  "The murine DSCR1-like gene family."
RL  Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE OF 40-239 FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX  MEDLINE=21085660; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ichii Y.,
RA  Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA  Flischman W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA  Sacchini L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Baren G.,
RA  Blake D., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA  Brownstein M.J., Bult C., Fletcher C., Fujita M., Gaitanari M.,
RA  Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaers P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA  Hayashaw-Borle A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA  Hayashizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection";
RJ  Nature 409:685-690(2001).
CC  -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC  by binding to the catalytic domain of calcineurin A. Could play a
CC  role during central nervous system development (by similarity).
CC  -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.emb-eb.ch/announce/
CC  or send an email to license@eb-sib.ch).
CC  EMBL: AF237888; AAF62539.1; -
DR  EMBL: AK019377; BAB31687.1; -
DR  MGD: MGI:1858220; Dscr112.
SQ  SEQUENCE 239 AA; 27153 MW; 1B2687B47B84D272 CRC64;

Query Match 63.7%; Score 663.5; DB 1; Length 239;
Best Local Similarity 72.0%; Pred. No. 9.3e-51;

```

		Query Match	63.3%; Score 658.5; DB 1; Length 241;
		Best Local Similarity	70.1%; Pred. No. 2.5e-50;
		Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;	
Qy	11	STLVACVDVEVTNQVKKEFGLFTYDECVTFQLFKSFRVRINFNSHPKSAARAE	70
Dd	47	TSLFACSVEAVFEAREKERFEALFTIYDDQVTFLFKSFRVRINFSPKPEAARAE	106
Qy	71	LHTQPRGKKLKYFAQVQTPTDGDKLHLAPOPAKOFLISPPSSPVGWKPISDATPV	130
Dd	107	LHETDFNGQKKLYFAQVOVSGEVRDKSYLLPQPVPKFLOISPASPVGWKOSEDAMP	166
Qy	131	LNVDLLAYAVAKLGPGKEYELHAGTSTPSVVHVCDSDMBEEDPKTSPPKILOTRRPG	190
Dd	167	INTDLLCNVSKLGPGKEYELHAGTSTPSVVHVCESETEEEBTK-NPKOKIAQTRRPD	225
Qy	191	LPPSVSN	197
Dd	226	PPTAAALN	232

[illegible]

```

DB 127 INVDLYAISKLGPSEKXELHAAITDTPSVVHVCSDDQENEEEMERMKRPKXIQT 186
QY 187 RRRGLPP 193
DB 187 RRRPEYTP 193

RESULT 6
CCPI_HUMAN
CCPI_HUMAN STANDARD; PRT; 197 AA.
AC P53805; O00583; O00583; Q9UPL5; Q9BU69;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calciopressin 1 (Down syndrome critical region protein 1) (Myocyte-
enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).
GN DSCR1 OR DSCR1 OR ADAPT78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RX MEDLINE=96121593; PubMed=8555418;
RA Fuentes J.-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
Estivill X.;
RT "A new human gene from the Down syndrome critical region encodes a
RT protein-rich protein highly expressed in fetal brain and heart.";
RL Hum. Mol. Genet. 4:1935-1944 (1995).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=97468152; PubMed=9325060;
RA Fuentes J.-J., Pritchard M.A., Estivill X.;
RT "Genomic organization, alternative splicing, and expression patterns
RT of the DSCR1 (Down syndrome candidate region 1) gene.";
RL Genomics 44:358-361 (1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
RX TISSUE=Breast;
RA Crawford D.R., Leahy K.P., Davies K.J.A.;
RT "Adapt78, a calcium and oxidant-inducible RNA.";
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20320698; PubMed=10861295;
RA Fuentes J.-J., Genesca L., Kingsbury T.J., Cunningham K.W.,
RA Perez-Riba M., Estivill X., de la Luna S.;
RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RT calcineurin-mediated signaling pathways.";
RL Hum. Mol. Genet. 9:1681-1690 (2000).
RN [6]
RP FUNCTION: Inhibits calcineurin-dependent transcriptional responses
RN by binding to the catalytic domain of calcineurin A. Could play a
RN role during central nervous system development.
RN [7]
RP ALTERNATIVE PRODUCTS: 4 ISOFORMS: 1 (SHOWN HERE), 2, 3 AND 4; ARE
RN PRODUCED BY ALTERNATIVE SPLICING.
RN [8]
RP TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
RN muscle. Also expressed in all other tissues.
RN [9]
RP INDUCTION: By calcium.
RN [10]
RP SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
RN [11]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.ebi.ac.uk/announce/
RN or send an email to license@ebi.ac.uk).

```

```

CC EMBL; U28833; AAB81557.1; -
DR EMBL; U85265; AAB84370.1; -
DR EMBL; U85266; AAB84371.2; -
DR EMBL; U85267; AAB84372.1; -
DR EMBL; BC002864; AAB02864.1; -
DR EMBL; U53821; AAF21218.1; -
DR GeneW; HGNC:3040; DSCR1.
DR MIM; 602917; -
CC Alternative splicing.
FT DOMAIN 166 174 POLY-GLU.
FT VASAPLIC 1 28 MEEVDLDLPATACHLDPVFDGLC -> MHPFRNPNYS
FT VASAPLIC 1 29 FSLIACVANDISEST (IN ISOFORM 2).
FT VASAPLIC 1 29 MEEVDLDLPATACHLDPVFDGLC -> MYV (IN
FT VASAPLIC 1 80 ISOFORM 3).
FT VASAPLIC 1 80 MISSING (IN ISOFORM 4).
FT CONFLICT 159 159 H -> Q (IN REF. 4).
SQ SEQUENCE 197 AA; 22637 MW; 1BD426BB8167BEC CRC64;

Query Match 59.8%; Score 622; DB 1; Length 197;
Best Local Similarity 63.6%; Pred. No. 3e-47;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLVAQVDVETNQNQVKEKEGELFRFYDECVTQQLFKSPRRVAINSHPSARARIE 70
DB 11 SYTIACHDPRFVFDGLCRANKESLFRYDKDITFOYKSPFRVRINSPSADARLQ 70
QY 71 LHETOPRGKLTLYPAQVOTPEIDGKHLAPPOPAKQFLSPSPVGMFPISDATPV 130
DB 71 LHKTEFLKEMKLYPAQTLHIQSS----HLAPNDKQFLSPSPVGMKQVDAITPV 126
QY 131 INVDLYIAVAKLGPSEKXELHAAITDTPSVVHVCSDDQENEEEMERMKRPKXIQT 186
DB 127 INVDLYIAISKLGPSEKXELHAAITDTPSVVHVCSDDQENEEEMERMKRPKXIQT 186
QY 187 RRRGLPP 193
DB 187 RRRPEYTP 193

RESULT 7
CCPI_MOUSE
CCPI_MOUSE STANDARD; PRT; 198 AA.
AC Q9UHG6; Q9UKX3; Q9UKX2; Q9UKS1; Q9UKS0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calciopressin 1 (Down syndrome critical region protein 1 homolog)
DE (Myocyte-enriched calcineurin interacting protein 1) (MCIP1).
GN DSCR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX STRAIN=BA1B/C; TISSUE=Brain;
RN MEDLINE=20534792; PubMed=11080588;
RA Striopol P., Petrini M., Lenzi L., Carinci P., Zamotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RT conserved syntenic with the human orthologous genes.";
RL Gene 257:223-232 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=20187590; PubMed=10722714;
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RT enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725 (2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.
RX TISSUE=Fetal brain;

```



Wang Z.-Y., Wasmann D.A., Weinrock G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 RL -1- FUNCTION: REQUIRED FOR THE DSCR1 FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF147700; AAD33987.1; -;  
 CC EMBL: AE003712; AAF55285.1; -;  
 CC FlyBase: FBgn0026629; nla.  
 DR PLYBASE: 292 AA; 31423 MW; 64F1BBF5FEAC6F9 CRC64;  
 SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5FEAC6F9 CRC64;  
 Query Match 38.8%; Score 404; DB 1; Length 292;  
 Best Local Similarity 43.2%; Pred. No. 4,2e-28;  
 Matches 83; Conservative 35; Mismatches 62; Indels 12; Gaps 4;  
 Oy 4 PSMDCD-----VSTLVACVDEVFNTNOEKEKEGFRFTYDECVTQFLQKRRRLNF 58  
 Db 104 PEVDADSDPDLPTSLITVNIHSEVFANPFLKNAMELFTFPESSATFQLRFRRLRVY 163  
 Oy 59 SHPSAARARIELHETQFGKK-LKLYPAQVQTPETDQKHLAPPAKQFLSPSPSP 117  
 Db 164 DNALAAANARILHGYEFNKKTIVTCYFAQPTVPSN---KNLGPAPVQKFLISPPASP 220  
 Oy 118 PVGMKPISDATVLYDLYAAKLGPGKXYELHAGTSPSVVAVHVCSDMEEDPRT 177  
 Db 221 PAGWEPREGELVNHDLAALSLTPGSHLHPOSEDPAIIVHTA---MLAETGPDL 277  
 Oy 178 SPKPIIOTRRP 189  
 Db 278 QVKAPIVOTKCP 289  
 RESULT 9  
 CCPL\_CAEEL STANDARD: PRT; 207 AA.  
 AC P53B06; Q8U6V5; 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Calcipressin-like protein (Down Syndrome candidate region 1-like protein)  
 GN RCN-1 OR DSCR1L OR P54E7.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bentley D.;  
 RL Submitted (JUN-1994) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA MEDLINE=20223370; PubMed=10756093;  
 RA Striopol P., Lenzi L., Petrini M., Carinci P., Zannotti M.;  
 RT "A new gene family including DSCR1 (Down Syndrome candidate region 1) and ZAL1-4: characterization from yeast to human and identification of DSCR1-like 2, a novel human member (DSCR1L2).";  
 RL Genomics 64:252-263(2000).  
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

by binding to the catalytic domain of calcineurin A (By similarity).  
 -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U00067; AK20076.1; -;  
 CC EMBL: AF176115; AAF01683.1; -;  
 CC WormPep; F54E7.7; CE01318.  
 DR CONFLICT 205 207  
 FT CEQ -> SLOGSSFLCTPSPSPFSPFPFPFPSPSPSPSP  
 FT G (IN REF. 1).  
 SQ SEQUENCE 207 AA; 23030 MW; 0154B308AB05B79 CRC64;  
 Query Match 31.0%; Score 322.5; DB 1; Length 207;  
 Best Local Similarity 38.9%; Pred. No. 3,4e-21;  
 Matches 68; Conservative 36; Mismatches 64; Indels 7; Gaps 4;  
 Oy 11 STLVACVDEVFNTNOEKEKEGFRFTYDECVTQFLKSPRRVINSHPKSAARARIE 70  
 Db 30 NAIITQVEDVDFDKQKANKFSSLTQIEKDIHPDLPKSRFRVRVIFSSPENATPAKLI 89  
 Oy 71 LHETQFGKKLKYPAQVQTPETDQKHLAPPAKQFLSPSPSPVGMKPISDATVPY 130  
 Db 90 VQGFSEKHELKAFPA--QRIYMSANSQMLSPPLEKQFLISPPSPVGMKPISDATVPY 147  
 Oy 131 L-NYDLYAAKLGPGKXYELHAGTSPSVVAVHVCSDMEEDPRTSPK 181  
 Db 148 VCNFPLMARLASFAIDEXEYVNGDELTPALIVHCPEDIVPAISLEMPRT-PRP 201  
 RESULT 10  
 YAGF\_SCHPO STANDARD: PRT; 163 AA.  
 ID YAGF\_SCHPO  
 AC 009751;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C13G6.15c in chromosome I.  
 GN SPAC13G6.15C OR SPAC24B11.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.D., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voicikart G., Aert R., Roben J., Grymopoulos B., Welter C., Vansirele E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B., Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaune V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

DR Genw; HGNC:16262; YAP1.
DR MIM; 606608; -.
DR InterPro; IPR005153; MbtH.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00357; WW; 1.
DR Pfam; PF03621; MbtH; 1.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 1.
DR KW Phosphorylation.
FT DOMAIN 171 204 WW
SQ SEQUENCE 454 AA; 48755 MW; 87CB840D3393BFC0 CRC64;

Query Match 9.7%; Score 101; DB 1; Length 454;
Best Local Similarity 32.1%; Pred. No. 0.16;
Matches 34; Conservative 10; Mismatches 30; Indels 32; Gaps

QY 102 PPOPAKQFLISPPSPVPGWKPISDATPVNLVDLLYAVAKLGGKYEHLHAGCTSTPSV- 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 PPOPAFGQGGQSPSPFPFGQGGPPS-----GFGQ--PAPAATQAAPQAP 48
QY 161 -----VVHV-CDSMEEEE-----DPKTSPPKPIQTRPGLPSS 194
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 PAGHIQVHVGRDSETDLALFNAMNPKTANPVQTPVPMRLKLPS 94

RESULT 12
CAP_YEAST STANDARD; PRT; 526 AA.
AC PL7555;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Adenylyl cyclase-associated protein (CAP).
GN SRV2 OR CAP1 OR YNL138W OR N1210 OR N1838.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90235281; PubMed=2184942;
RA Field J., Vojtek A., Ballester R., Bolger G., Colicelli J.,
RA Ferguson K., Gerst J., Kataoka T., Michaeli T., Powers S., Riggs M.,
RA Rodgers L., Wieland I., Wheland B., Wigler M.;
RT "Cloning and characterization of CAP, the S. cerevisiae gene encoding
RT the 70 kd adenylyl cyclase-associated protein.";
RL Cell 61:319-327(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=90235282; PubMed=2159860;
RA Fedor-Chaikin M., Deschenes R.J., Broach J.R.;
RT "SRV2, a gene required for RAS activation of adenylyl cyclase in
RT yeast.";
RL Cell 61:329-340(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=96109932; PubMed=8619318;
RA Mallet L., Bussiere F., Jacquet M.;
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,
RT MEP2, CAP/SRV2, NAM5, FKBI/FPRI/RBPI, MOM22 and CPT1, predicts an
RT adenosine deaminase gene and 14 new open reading frames.";
RL Yeast 11:1195-1209(1995).
CC -!- FUNCTION: THE N-TERMINAL DOMAIN BINDS TO ADENYLYL CYCLASE, THEREBY
CC ENABLING ADENYLYL CYCLASE TO BE ACTIVATED BY UPSTREAM REGULATORY
CC SIGNALS, SUCH AS RAS. THE C-TERMINAL DOMAIN IS REQUIRED FOR NORMAL
CC CELLULAR MORPHOLOGY AND GROWTH CONTROL.
CC -!- SUBCELLULAR LOCATION: CELL MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CAP FAMILY.

```







```

RA MEDLINE=88274407; PubMed=2455776;
RX Kosik K.S., Orcicchio L.D., Bakalis S., Duffy L., Neve R.L.;
RT "Partial sequence of MAP2 in the region of a shared epitope with
RL Alzheimer neurofibrillary tangles.";
CC J. Neurochem. 51:587-598(1988).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U01828; AAA03354.1; -
CC EMBL; U89330; AAB48098.1; -
CC EMBL; U89329; AAB48097.1; -
CC EMBL; M25668; AAA59552.1; -
CC PIR; PL0024; QRHUM1.
CC Genew; HGNC:6839; MAP2.
CC MIM; 157130; -
CC InterPro; IPR001084; Tubulin Tau.
CC Pfam; PF00418; tubulin-binding; 3.
CC PROSITE; PS00329; TAU_MAP; 2.
KW Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
FT DOMAIN 1447 1467 CALMODULIN-BINDING (POTENTIAL).
FT REPEAT 1661 1691 TAU/MAP MOTIF.
FT REPEAT 1692 1722 TAU/MAP MOTIF.
FT REPEAT 1723 1754 TAU/MAP MOTIF.
FT VARSPLIC 152 1507 MISSING (IN ISOFORM MAP2C).
FT CONFLICT 9 9 A -> G (IN REF. 2).
FT CONFLICT 37 37 R -> A (IN REF. 2).
FT CONFLICT 108 108 A -> G (IN REF. 2).
FT CONFLICT 152 155 MISSING (IN REF. 2).
FT CONFLICT 187 187 S -> K (IN REF. 2).
FT CONFLICT 1655 1655 A -> GL (IN REF. 2).
FT CONFLICT 1736 1736 A -> A (IN REF. 2).
SQ SEQUENCE 1827 AA; 199610 MW; BAC36D0030F5F455 CRC64;

Query Match 8.1%; Score 84.5; DB 1; Length 1827;
Best Local Similarity 35.9%; Pred. No. 23;
Matches 28; Conservative 6; Mismatches 31; Indels 13; Gaps 3;

QY 101 APPQAKQFLISPPSPVGVGWKPISDATPVLYNLYAVAKLGPGEKVELHAGTSTPSV 160
Db 122 ALPLAAETANLPPSPPP---SPASEQTIVVEDLLTA-----SKMEPHDQQLTPST 171

QY 161 VHVCDSDMEEEEDPKTS 178
Db 172 AE---PSDQKEKESKQS 186

RESULT 15
FM14_MOUSE
ID FM14_MOUSE STANDARD; PRT; 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Formin 1 isoform IV (limb deformity protein).
GN FMN OR LD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
CC Grusby-Jackson L., Kuo A., Leder P.;
RT "A variant limb deformity transcript expressed in the embryonic mouse
RL limb defines a novel formin.";
CC Genes Dev. 6:29-37(1992).
CC -!- FUNCTION: IS IMPORTANT IN THE MORPHOGENESIS OF LIMB AND MAY HAVE A
CC FUNCTION IN DIFFERENTIATED CELLS OR BE INVOLVED IN MAINTAINING
CC SPECIFIC DIFFERENTIATED STATES.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: IA (AC Q05860), IB (AC
CC Q05860), II (AC Q05860), III (AC Q05860) AND IV (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING. A VARIATION IN SPLICING IS SEEN
CC AMONG DIFFERENT TISSUES AND DIFFERENT SIZE TRANSCRIPTS EXIST
CC WITHIN ANY ONE TISSUE.
CC -!- TISSUE SPECIFICITY: IT IS FOUND THROUGHOUT THE EMBRYO BUT
CC HAS A FUNCTIONAL ROLE ONLY IN THE KIDNEY AND LIMB.
CC -!- DEVELOPMENTAL STAGE: THIS IS THE ISOFORM FOUND IN THE APICAL
CC ECTODERMAL RIDGE AND THE MESENCHYMAL COMPARTMENT OF THE DEVELOPING
CC LIMB BUD.
CC -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 1 (FH1) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 2 (FH2) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE FORMIN HOMOMOLOGY FAMILY. CAPPUCCINO
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62379; CAM44244.1; -
CC PIR; S24407; S24407.
CC MGD; MGI:101815; Fmn.
CC InterPro; IPR003104; FH2.
CC InterPro; IPR01265; Formin.
CC Pfam; PF02181; FH2; 1.
CC PRINTS; PR00828; FORMIN.
CC SMART; SM00498; FH2; 1.
KW Nuclear protein; Developmental protein; Alternative splicing;
KW Phosphorylation; Coiled coil.
FT DOMAIN 418 443 COILED COIL (POTENTIAL).
FT DOMAIN 457 566 COILED COIL (POTENTIAL).
FT DOMAIN 644 744 FH1 (PRO-RICH).
FT DOMAIN 759 1164 FH2.
FT DOMAIN 1043 1116 COILED COIL (POTENTIAL).
FT DOMAIN 635 638 POLY-SER.
FT DOMAIN 751 755 POLY-SER.
SQ SEQUENCE 1206 AA; 133464 MW; 4DFB38CB52BD8EE7 CRC64;

Query Match 8.0%; Score 83; DB 1; Length 1206;
Best Local Similarity 23.4%; Pred. No. 19;
Matches 46; Conservative 18; Mismatches 87; Indels 46; Gaps 8;

QY 26 QEVKEKFEGLFRITYDECVTFQ---LFKSPRV-----RINFSPKSAARARIELHETQFR 77
Db 547 QQLEKRECEEMRDVCISTDDDCSPKAFNVCIQDTRETFKPCDAESK--ATRSQIV 604

QY 78 GKKLKLYFAQVQTPETDGDGLHL-----AP---PQPAKQFLISPPSS 116
Db 605 PKLLITSLTQL-SPKSDKDIHAPFQTRGTSSSSQOKISPPAPPTPPPLIPPLIPPPP 663

QY 117 PPVGMKPISDATPVLYNLYAVAKLGPGEKVELHAGTSTPSVVVHVCDSDMEEEEDPK 176
Db 664 LPPGLGLPLPPAPPI-----PPVCPVSP-----PPPPPPPPPTVPVPSDGPDPPPP 709

QY 177 TSPKPKIITRPPGPP 193
Db 710 PPPLPNVLALPNSGGPP 726

```

Sun Dec 15 08:39:01 2002

us-09-782-953-9.rsp

Page 10

Search completed: December 11, 2002, 11:39:43  
Job time : 9.3086 secs

---

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:40 ; Search time 12.9562 Seconds  
(without alignments)  
1461.736 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCDVSTLVACVVDV.....SPKPKIIQTRRGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	322.5	31.0	239	T34305	hypothetical prote
2	133.5	12.8	315	T49641	nebula related pro
3	107.5	10.3	163	S62444	conserved hypothet
4	101	9.7	454	A54954	yes-associated pro
5	93	8.9	1018	JC4211	neural adhesion pr
6	92	8.8	277	T37629	hypothetical prote
7	90	8.6	526	A34896	adenylate cyclase-
8	87.5	8.4	211	S37792	hypothetical prote
9	87.5	8.4	13055	T16580	hypothetical prote
10	87	8.4	1018	T19693	hypothetical prote
11	86	8.3	1006	T42731	hypothetical prote
12	85.5	8.2	7576	T17428	atrophin-1 related
13	84.5	8.1	212	T48580	FK506 polyketide s
14	84.5	8.1	1222	B90593	hypothetical prote
15	84.5	8.1	3938	T42761	hypothetical prote
16	84	8.1	671	T29634	Basoon protein--
17	84	8.1	1557	T13160	probable Poly-A Bi
18	83	8.0	160	C72706	protein CNK - frui
19	83	8.0	643	T23453	hypothetical prote
20	83	8.0	746	S74546	glycogen operon pr
21	83	8.0	1206	S24407	formin isoform IV
22	83	8.0	1322	S40766	hypothetical prote
23	83	8.0	1468	G11515	formin - mouse
24	82.5	7.9	656	G86482	protein F5J5.9 [im
25	82	7.9	1611	T38236	hypothetical prote
26	82	7.9	2464	1 ORMSP1	microtubule-associ
27	82	7.9	2783	1 A41948	alpha-fetoprotein
28	81	7.8	559	2 A57474	extracellular matr
29	81	7.8	813	2 A72203	cellobiose-phospho

30 81 7.8 2364 2 A56577 microtubule-associ  
31 80.5 7.7 1015 2 JC6552 DNA topoisomerase  
32 80 7.7 863 2 S06017 neuraxin - rat  
33 79.5 7.6 3942 2 T42730 Basoon protein -  
34 79.5 7.6 6831 2 A88852 protein unc-22 [im  
35 79.5 7.6 6839 2 S57242 twitchin [similar  
36 79.5 7.6 7160 2 T27935 hypothetical prote  
37 79 7.6 909 2 S32538 CGMP-gated cation  
38 79 7.6 992 2 T46337 hypothetical prote  
39 79 7.6 1211 2 T42230 AF4 protein - mous  
40 78.5 7.5 346 2 I48185 gene alx3 protein  
41 78.5 7.5 497 2 S22708 homeotic protein e  
42 78.5 7.5 737 2 S28030 DNA-binding protei  
43 78.5 7.5 836 2 S49940 cell division cont  
44 78 7.5 327 2 D72503 probable modificat  
45 78 7.5 1018 2 A54744 contractin 1 precur

#### ALIGNMENTS

##### RESULT 1

T34305  
hypothetical protein F54E7.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34305

R:Bentley, D.

Submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid F54E7.

A:Reference number: Z21502

A:Accession: T34305

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-239 <BEN>

A:Cross-references: EMBL:U00067; PIDN:AAC77519.1; GSPDB:GN00021; CESP:F54E7.7

A:Experimental source: strain Bristol N2; clone F54E7

C:Genetics:

A:Gene: CESP:F54E7.7

A:Map position: 3

A:Introns: 40/3; 78/2; 165/2; 205/1

Query Match 31.0%; Score 322.5; DB 2; Length 239;

Best Local Similarity 38.9%; Pred. No. 3.6e-22;

Matches 68; Conservative 36; Mismatches 64; Indels 7; Gaps 4;

QY 11 STLIVACVVDVEFTNQSVKEKEGLRPTDYDECVTFLPKSPRRVRINFSPKSAARIE 70

Db 30 NAIITVQVPEDVFDNKNQKANKFSSLTQIEKDIHFDLRSFRVRVIFSPENATAAKLI 89

QY 71 LHETQFRGKKLYFAQVOTPETDGDKLHLPQAKQFLISPPSPPPVGVWKPISDATPV 130

Db 90 VQGFSEKHELKAFPA--ORIYMSANSQMLSPPLEKQFLISPPSPPPVGVWQTKMPVP 147

QY 131 L-NYDLLYAVAKLPGKEKVELHAGTESTPSVVHVCDSDMEEE---EDPKTSKP 181

Db 148 VCNFDLMARLASFAIDKEKVEVNGDELTPAIIVHPCETIDVPSAEMPT-PRP 201

##### RESULT 2

T49641

nebula related protein [imported] - Neurospora crassa

N:Alternate names: protein B5022.230

C:Species: Neurospora crassa

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49641

R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Nyakatura

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49641

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <SCH>



Db 913 SSVRSGSRYITWDHVVALSNESVTGKYKLYRPDQGDGKLYSTHKHSIEVP 965

## RESULT 6

hypothetical protein SPAC13F5.04c - fission yeast (*Schizosaccharomyces pombe*)

C;Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 03-Dec-1999

C;Accession: T37629

R; Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream M.A.; Wood V.

submitted to the EMBL Data Library September 1997

**A:Reference number:**

A: Accession: T37629  
R: Reference Number: 421733

A/ACCESSION: 137629

A; status: preliminary

A; Molecule type: DNA

A;Residues: 1-277 <BRO>

A;Cross-references: EMBL:Z99091; PIDN:CAB11767.1;

A; Experimental source: ST

C:Genetica:

A:Gene: SPDB.SPAC

A:Gelle: S&UB:SP:  
A:Man position:

Query Match 8.8; Score 92; DB 2; Length 277;  
Best Local Similarity 22.0; Pred. No. 0.56;  
Matches 45; Conservative 24; Mismatches 76; Indels 60; Gaps 7;

0V 42 CVTEOLFKEFRRVRINE-----SHDYCAZB-----ARTER HEMORRHOY 20

27 CVTFQFASFKRVRI-----SHPKSAAR-----ARIELHETQFRGK 79

[illegible]

DB 2 CFFRLVSSPAILRISFLSITKSIKFKLRLQVANPQRKAKTPSNHATEELQQSSNST 61

1

QY 80 KKLKLYFAQVQTPETDGDGDKLHLAPP--QPAKQFLISPPSSPPVGVKPP-----ISDA 12

[illegible]

db 62 TLPTOEAAVETNASASASHTSFALPTTSPASISLSTPSTKSAIVSSCFNNVFRNDKWSI CCFNN 12

```
Qy 128 TPLVNYDLLYAVAKLGPEKYE-----LHAQTSTSPV-----VVHVCDSDMEE 171
Db 122 APQLN-----SFSHSVEPTFFSTTSITENLTIDPTRSTRSSSHIQSLSPSEKQ 171
Qy 172 EEDPKTSPKPKIIOTRRRGLPPSPVS 196
Db 172 TSDGHRPPSPSTITTTSIDPSVA 196
```

RESULT 7  
A344896

A:Molecule type: DNA  
A:Residues: 1-526 <MAL>  
A:Cross-references: EMBL:Z46843; NID:g861113; PIDN:CAA86887.1; PID:g854499  
R:Mallet, L.; Bussereau, F.; Jacquet, M.  
A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MPA2, MEP2, CAP/SRV2  
A:Reference number: S59241; MUID:96109932; PMID:8619318  
A:Accession: S59250  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-526 <MAW>  
A:Cross-references: EMBL:Z46843; NID:g861113; PIDN:CAA86887.1; PID:g854499  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994  
R:Mallet, L.; Bussereau, F.; Jacquet, M.  
A:Submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63069  
A:Accession: S63083  
A:Molecule type: DNA  
A:Residues: 1-526 <MAP>  
A:Cross-references: EMBL:Z71144; NID:g1302083; PIDN:CAA96020.1; PID:g1302084; MIPS:YNL1  
A:Experimental source: strain S288C

A: Map position: 14L  
A: Cross-references: SGD:S0005082; MIPS:YNL138W  
C: Superfamily: adenyl cyclase-associated protein MCH1

Query Match 8.6%; Score 90; DB 2; Length 526;

Best Local Similarity 24.9%; Pred. No. 2.3;

Matches	52;	Conservative	26;	Mismatches	97;	Indels	34;	Gaps	34;
---------	-----	--------------	-----	------------	-----	--------	-----	------	-----

[illegible]

QV 10 VSTI.VACWU - - - - VEVETNOIVKEKECEI.EBTVDECIITEOI.KKCEBBRITVKECHNVCAN.CC

QY VSILVACVVD---VEVFINQEVKREFGLEFRTYDECVTFLQLFKSFRRVRINFSHPKSAA 65

[illegible]

```

Qy 66 PARLELHETQPRGKKLYIFYAQVOTPTDGDKHLHAPQPAKQFLI-----SPPSSPPVGM 121
Db 233 K----EYHTTGVSWKKDGMDFADAMAQSTTKMTGA--TSSPSASATAAPAPPPPPAPPASV 288
Qy 122 KPISDAPVLNLYDL----LVAV-AKLGPGE-----KYELHAGTESTPSVVVHVCDSDM 169
Db 289 FEISNDTPATSSDANKGGIGNVFAELNOGENITKGLKKVDKASQOQTHKNPELRQSTVSST 348
Qy 170 BEEDPKTSP-KPKIIOTRRRGLPPSVSN 197
Db 349 GSKSGPPPPPKPKSTLTKRPPRKELVGN 377

RESULT 8
S37792
hypothetical protein YKULS9c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YKL613
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence revision 03-May-1994 #text_change 19-Apr-2002
C/Accession: S37792; S37989; 544569
R/Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL data Library, September 1993
A/Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci
A/Reference number: S37786
A/Accession: S37792
A/Molecule type: DNA
A/Residues: 1-211 <VAN>
A/Cross-references: EMBL:Z26877; NID:g407482; PIDN:CAA81495.1; PID:g407489
A/Experimental source: strain S288C
R/Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994

```

Yeast 10, 35-40, 1994

A:Title: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loci of

A:Reference number: S44563

A:Accession: S44569

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-211 &lt;VA3&gt;

A:Cross-references: EMBL:226877; NID:9407482; PIDN:CAA01495.1; PID:9407489

A:Experimental source: strain S288C

A&gt;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993

C:Genetics:

A:Gene: SGD:RCN1

A:Cross-references: SGD:S0001642

A:Map position: 11L

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YKL159C

Query Match 8.4%; Score 87.5; DB 2; Length 211;

Best Local Similarity 25.2%; Pred. No. 1.2; Indels 31; Gaps 6;

Matches 31; Conservative 23; Mismatches 38; Indels 31; Gaps 6;

QY 8 CDVSTLVACVD-----VEFTNQVEKEFEGLFRTYDECVTFFQKSFRRVRI----- 56

DB 16 CD-----IVDNDNERIGVMISKNILRFG---INENBPILQILILKKFKRILILICPSH 65

QY 57 -NESHKSAARARIELHETQFRGKULKYFAQVQPTETDGLHLAPQPAKQFLISPS 115

DB 66 DISOHVMDASRA-LEMENFNFS-----YSLQDQORNLTKQYLKVPSEKMFILISPA 116

QY 116 SPP 118

DB 117 SPP 119

## RESULT 9

T16580 hypothetical protein K07E12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16580

R:Fullon, L. submitted to the EMBL Data Library, May 1994

A:Description: The sequence of C. elegans coemid K07E12.

A:Reference number: Z18540

A:Accession: T16580

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1305 &lt;PIL&gt;

A:Cross-references: EMBL:U00054; NID:9485140; PID:9485141; PIDN:AAA50715.1; CESP:K07E12.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K07E12.1

A:Introns: 46/3; 85/2; 201/1; 278/2; 470/2; 817/1; 960/3; 1265/1; 1332/1; 1478/3; 6014/3; 6159/3; 6665/2; 7266/3; 7895/3; 8669/3; 8726/3; 9803/3; 10937/3; 12234/2; 1229

Query Match 8.4%; Score 87.5; DB 2; Length 1305;

Best Local Similarity 23.4%; Pred. No. 2.2e+02;

Matches 54; Conservative 27; Mismatches 179; Indels 71; Gaps 11;

QY 1 MPAPSMDCDSTLVACVDVEFTN-----OEVEKEFGLFR-TYDECVT-FQL 47

DB 447 MTAFEDICNTR--ADPVPLVWLHKRPLNKGSKTOHKKKGCVLESTQFSCVAENEA 504

QY 48 FKSFRRIRINFSHPKSAARARIELHETQFRGKULKYFAQVQPTETDGLHLAP--APQP 105

DB 505 GKSTKILNVTGSAERIRYQI-----DGDVTLQWMEPQI 542

QY 106 AK-----OFLISPPSPVGVK-----PISDAPVLNYD-----LLYAVAKLGP 145

DB 543 TNGPMAGYDVYETDPSLPRQWKHHIDDPNATTIVLRLNKKTPYTFVIYVGNRLPG 602

QY 146 ---EKYLHACTESTPSVYVAVCDSDMEBED-----PKTSPPKII 184

DB 603 LPSPAFATITWLAKPPVQLEPSEMTKEPSNDEMIIECGAGCVKPKII 653

## RESULT 10

T19693 hypothetical protein C34B7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T19693

R:Harris, B.

submitted to the EMBL Data Library, December 1996

A:Reference number: 219165

A:Accession: T19693

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1018 &lt;WIL&gt;

A:Cross-references: EMBL:283220; PIDN:CA05700.1; GSPDB:GN00019; CESP:C34B7.1

A:Experimental source: clone C34B7

C:Genetics:

A:Gene: CESP:C34B7.1

A:Map position: 1

A:Introns: 78/1; 149/1; 177/1; 245/1; 271/1; 311/2; 361/2; 387/3; 415/2; 444/3; 486/2; 5

Query Match 8.4%; Score 87; DB 2; Length 1018;

Best Local Similarity 24.0%; Pred. No. 9.8;

Matches 46; Conservative 23; Mismatches 85; Indels 38; Gaps 8;

QY 1 MPAPSMDCDV-----TLVACVDVEFTNQVEKEFEGLFRTYDECVTFFQKFS 50

DB 337 LPTAPQNSVSSKVCNABILITLFTKVEKSVSPSLNPEPLESLFRNPANCFENIPLYL 396

QY 51 F--RRVIRINFSHPKSAARARIELHETQFRGKULKYFAQV---QTPETDDK-LHLAP 103

DB 397 IFNRKMGKSTKPTSKANSSPE-----GRAMSPKQTLLDDQTPSSSSSKKNLKKP 449

QY 104 QPAKQFLISPPSPVGVKPISDATPVNLVLYAVAKLGPGEKYLHAGTESTPSVYVH 163

DB 450 STDD-----TPSSSEPLSKNRLDSTETAEED-----NTRRRRTLKGVALP--VCS 495

QY 164 VCDSDMEBEDP 175

DB 496 FCKEDQKADDEP 507

## RESULT 11

T42731 atrophin-1 related protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C:Accession: T42731

R:Khan, F.A.; Margolis, R.L.; Loev, S.L.; Sharp, A.H.; Li, S.H.; Rose, C.A.

submitted to the EMBL Data Library, December 1995

A:Description: CDNA sequence and expression of an atrophin-1 (DRPLA disease gene) relate

A:Reference number: Z22250

A:Accession: T42731

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1006 &lt;KIA&gt;

A:Cross-references: EMBL:U44091; NID:91297310; PID:91209103; PIDN:AAA98970.1

C:Genetics:

A:Gene: ARP

Query Match 8.3%; Score 86; DB 2; Length 1006;

Best Local Similarity 28.2%; Pred. No. 12;

Matches 33; Conservative 13; Mismatches 43; Indels 28; Gaps 7;

QY 91 PETDGDKLHAP---PQAKQFL-----ISPPSPVGVKPI---SDATPVNLVLYA 138

DB 459 PTTGG--LHQVDSQSPFPQHPFVPGCPPTPTPSCPTPTSPAGSSSSQPCSS----- 510

QY 139 VAKLGPGEKYLHAGTESTPSVYVAVCDSDMEBEDPKTSPKXIIQTRRPLGPSV 195

DB 511 -AAVSGGNV---PGAPSCPLPAVQIKEDALDEABEPBPFP---PRSPBPFTV 559

## RESULT 12

T17428  
FK506 polyketide synthase - Streptomyces sp. (strain MA6548)  
C:Species: Streptomyces sp.  
A:Variety: strain MA6548  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000  
C:Accession: T17428  
R:Motamedi, H.; Shafiee, A.  
Eur. J. Biochem. 256, 528-534, 1998  
A:Title: The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506 polyketide synthase  
A:Reference number: Z18779; MUID:98451508; PMID:9780228  
A:Accession: T17428  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-7576 <MOT>  
A:Cross-references: EMBL:AF082100; NID:G3798623; PID:G3798624; PIDN:AAC68815.1  
A:Experimental source: strain MA6548  
C:Genetics:  
A:Gene: ftkB  
C:Function:

A:Description: involved in synthesis of the backbone of the immunosuppressant FK506 polyketide  
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase  
C:Keywords: carrier protein  
F:54-500/Domain: acetate-CoA ligase homology <ACL>  
F:1095-1166/Domain: acyl carrier protein homology <ACP1>  
F:1204-1599/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
F:1682-1953/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>  
F:2680-2751/Domain: acyl carrier protein homology <ACP2>  
F:2804-3198/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>  
F:3295-3569/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>  
F:4320-4391/Domain: acyl carrier protein homology <ACP3>  
F:4435-4830/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>  
F:5903-5974/Domain: acyl carrier protein homology <ACP4>  
F:6018-6412/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>  
F:6513-6785/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>  
F:7450-7521/Domain: acyl carrier protein homology <ACP5>

Query Match 8.2%; Score 85.5; DB 2; Length 7576;  
Best Local Similarity 26.9%; Pred. No. 1.7e+02;  
Matches 39; Conservative 17; Mismatches 60; Indels 29; Gaps 7;

QY 47 LFPSRRVRINFS-----HPKSAARARIELHETQRGKKLYFAQVQ--TPPTDGDKLH 99  
DB 3834 LFADPERRLAFWSGVRHARAAATQORVLLRAD--GDTIRILATDERHGAFLVDYDGLT 3891  
QY 100 LAPPQPAKQEL-----ISPPSPPVGVKPKTSDATPVLYNVDLLYAVAKLGPGEKVELHAGT 154  
DB 3892 VRAAEPTQALFVAVMVSPASVPVGMTHLAD---VPEGDVL-----EGDAGEGDVGE 3941  
QY 155 ESTPSVVVHVCDSDMEEDPKTSP 179  
DB 3942 SDPPPPVVVLAV-----EPGDPDSSP 3961

## RESULT 13

T48580  
hypothetical protein T31B5.150 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48580  
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24490  
A:Accession: T48580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <BEV>  
A:Cross-references: EMBL:AL163491  
A:Experimental source: cultivar Columbia; BAC clone T31B5  
C:Genetics:  
A:Map position: 5  
A:Introns: 32/1

A:Note: T31B5.150

Query Match 8.1%; Score 84.5; DB 2; Length 212;  
Best Local Similarity 26.9%; Pred. No. 2.3;  
Matches 35; Conservative 14; Mismatches 38; Indels 43; Gaps 7;  
QY 61 PKSAARARIELHET-----QPRGKKLYFAQVQ--VOTPTD-----GD 96  
DB 57 PKKAARVVLGTFTAEAAALAYDRAALKFKGTAKLNFFERVQGPPTTTTISHAPRGVSE 116  
QY 97 KULHAPPQPAKQFLISPPSPPVGVKPKTSDATPVLYNVDLLYAVAKLGPGEKVELHAGTE- 155  
DB 117 SMNSPPPRP-----GPPSTTTTSW-PMT-----YNQDILQYLAQLTNSNEVDLSYVTST 164  
QY 156 -----STPS 159  
DB 165 LPSQPFSTPS 174

## RESULT 14

B90593  
hypothetical protein MYPV 6500 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: B90593  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: B90593  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1222 <KUR>  
A:Cross-references: GB:AL445566; PID:gl4090065; PIDN:CAC13823.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPV 6500  
A:Genetic code: SGC3

Query Match 8.1%; Score 84.5; DB 2; Length 1222;  
Best Local Similarity 22.0%; Pred. No. 21;  
Matches 41; Conservative 34; Mismatches 68; Indels 43; Gaps 8;  
QY 9 DVSTLVACVDDVEFTNOEVKEKFEGL---FRTYDECVTQFLFKSFRVRINFSPKSA 65  
DB 218 DVTASAKIRVNAKLTNNINNKFEVINGFKY--VIADSLKAKFTLVK-----KDIA 269  
QY 66 RARIELHETO-FRGKKLYFAQVQTPETDGDKLHAPPQPAKQFLISPPSPPVGVKPI 124  
DB 270 IBEKRVVTEAFKSKTQEFASLK-----NYLDISLPEGLDPSFVS-----IETKND 317  
QY 125 SDATPVLYNVDLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDME-----EEDDP 175  
DB 318 DPESAILTYKLVK-----NLHSGSQANPSKNGVSVESKLKTFEIKRSQIQESQP 367  
QY 176 KTSKPK 181  
DB 368 KPCKPK 373

## RESULT 15

T42761  
Bassoon protein - rat  
N:Alternate names: brain-specific synapse-associated protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T42761  
R:Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex, J. Cell Biol. 142, 499-509, 1998  
A:Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized in the presynaptic active zone  
A:Reference number: Z22249; MUID:98345363; PMID:9679147  
A:Accession: T42761  
A:Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA  
A: Residues: 1-3938 <DIE>  
A: Cross-references: EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G3413504  
A: Experimental source: ectrain Sprague Dawley; brain  
C: Function:  
A: Description: may be involved in cytomatrix organization at the site of neurotransmitter  
A: Note: component of the presynaptic cytoskeleton  
C: Keywords: coiled coil; zinc finger

```

Query Match 8.1%; Score 84.5; DB 2; Length 3938;
Best Local Similarity 29.8%; Pred. NO. 91;
Matches 34; Conservative 17; Mismatches 36; Indels 27; Gaps 8

QY 86 AQCQPEPDGDKHLAPAPAKOFLSPSSPEVQWKPLSDAPVLYNLLVAVAKIGG 145
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1441 ASREKRLSGDCG-EVQPPQPSKGYSTFGSSP-----PLSPSP--SESPTEPSKLCGR 1492
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 146 EKXELHAGES-TPSVVHVHCSDMEEDDKPT--SPKPKIIQ---TRPGIL 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1493 ATARESTQTPSLTPS-----SDI-----PRVGVSPRPMVAGGTQPHRPSPT 1534
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: December 11, 2002, 11:38:40  
Job time : 19.9562 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1647.42 Seconds  
(without alignments)  
3480.130 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MFAPSMDCVSTLVACVVDV.....SPKPKIQRTRRGLPPSVSN 197

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh  
-Q/cgn2 1/USPTO spoof/US09782953/runat 11122002 114429 17340/app query.fasta\_1.1173  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=i -END=-1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09782953 @CNC 1.1.2844 @runat 11122002 114429 17340 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DSV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sas.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_ma.\*  
21: em\_ox.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pi.\*  
26: em\_ro.\*  
27: em\_sas.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_hugo\_hum.\*  
40: em\_hugo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	594	6	AX365318 Sequence
2	1041	100.0	594	10	AF237791 Mus muscu
3	1041	100.0	659	10	AF237887 Mus muscu
4	1027	98.7	3227	10	AB061524 Mus muscu
5	1007	96.7	3184	6	AX329596 Sequence
6	1007	96.7	3184	6	AX333015 Sequence
7	1007	96.7	3184	6	AX365327 Sequence
8	1007	96.7	3184	6	AX420436 Sequence
9	1007	96.7	3184	9	HUMZAKI4
10	926	89.0	3295	10	AB061525 Mus muscu
11	910	87.4	1021	6	AX420427 Sequence
12	905	86.9	934	6	AX074351 Sequence
13	905	86.9	3159	6	AX420425 Sequence
14	905	86.9	3261	9	AK090990 Homo sapi
15	892	85.7	3240	9	AY034085 Homo sapi
16	892	85.7	3253	9	AY034086 Homo sapi
17	663.5	63.7	776	10	AF237888 Mus muscu
18	658.5	63.3	828	6	AX365330 Sequence
19	658.5	63.3	828	6	AX420437 Sequence
20	658.5	63.3	828	9	AF176116 Homo sapi
21	643	61.8	626	10	AB075973 Rattus no
22	641	61.6	2216	10	CGU60263 Cricetus
23	641	61.6	2331	6	AX365321 Sequence
24	641	61.6	2346	9	HSU8267 Homo sapien
25	634.5	61.0	2141	10	AF263239 Mus muscu
26	634.5	61.0	2198	10	BC013551 Mus muscu
27	633.5	60.9	720	6	AX365333 Sequence
28	633.5	60.9	720	9	AF176117 Homo sapi
29	632.5	60.8	597	6	AX365315 Sequence
30	632.5	60.8	597	10	AF237790 Mus muscu
31	628	60.3	2348	6	AX281651 Sequence
32	626.5	60.2	2212	6	AX365324 Sequence
33	626.5	60.2	2227	9	HSU85266 Homo sapien
34	626.5	60.2	2289	9	BC002864 Homo sapi
35	626.5	60.2	2407	9	AK092184 Homo sapi
36	615.5	59.1	621	10	AF282255 Mus muscu
37	615.5	59.1	2125	10	AF260717 Mus muscu
38	613.5	58.9	597	6	AX365312 Sequence
39	613.5	58.9	597	10	AF237789 Mus muscu
40	601	57.7	2173	6	AX410694 Sequence
41	601	57.7	2173	9	HSU28833 Homo sapien
42	600	57.6	2174	6	AR034241 Sequence
43	598	57.4	2284	9	HSU85265 Homo sapien
44	595	57.2	615	6	AX420432 Sequence
45	593	57.0	798	9	AF400429 Homo sapi

# ALIGNMENTS

RESULT 1

AX365318		594 bp	DNA	linear	PAT 15-FEB-2002
LOCUS	AX365318				
DEFINITION	Sequence B from Patent WO0204491.				
ACCESSION	AX365318				
VERSION	AX365318.1	GI:18697047			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
JOURNAL	Williams, S.R. and Rocheme, J.B. Methods and compositions relating to muscle selective calcineurin interacting protein (mcip) Patent: WO 0204491-A 8 17-JAN-2002; Board of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rocheme, Beverly (US)				
FEATURES	Location/Qualifiers				
Source	1..594				
CDS	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	1..594				
	/note="unamed protein product"				
	/codon_start=1				
	/protein_id="CAD23810.1"				
	/db_xref="gi:18697048"				
	/translation="MPAPSMDCVSTLVACVDVEFTNQEVKEFEGLPTTYDECVT FQLFKSFRRVRINFSPKSAAR,ELHETDPRGKKLIFYAQVTPEITDGKLALALA POPAPFLISPPSSPVMKPISDATPVNLDLYAVANKYGEYEHLAGTESPPS" VVHVQDSMEEREDPKTPKRIOTRBPGLRPVSNN"				
BASE COUNT	148 a 165 c 149 g 122 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,49e-84	Length:	594		
Score:	1041.00	Matches:	197		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Gaps:	0		
DB:	6	Indels:	0		
US-09-782-953-9 (1-197) x AX365318 (1-594)					
OY	1 MEFPRAlAProSeRteRrArPyvAsApvAlseThrLeuValAlAcvValAlAspVal	20			
Db	1 ATGCCAGCCCCCTAGCACTGGACTGTGAAGTTGCCACTTGCGCTGTGTGCATGTC	60			
OY	21 GIUVAIphethrfaAngInguVallyvgJulyBPhgQuJcyldeurheargThTyraSp	40			
Db	61 GAGGCTCTTAACAATCAGAGGAGTTAAAGAAAATTGAGGAGCATGTTCCGCACTATGAT	120			
OY	41 GIUCvValThrPhgInleuPhelySerPharqArgValArgIIeaAnPheserHis	60			
Db	121 GAATGTGTGTCGTTCCAGCTGTTTAAGAGTTTCCAGGGGTTTGCAMAATTTTCAGCAT	180			
OY	61 ProlvSerAlalAlarqAlaArgIIegJuleuHieglurhgInPhearsgJLylys	80			
Db	181 CCCAAATCTGCACCCCGTGCCCGGATAGAAGCTTCAATAGACTCAGTTCAAGGAGAG	240			
OY	81 LeuLyLeuTyrrheaIagInvalginThrProgluThraPgIyaBplysLeuHiblen	100			
Db	241 CTAAAACTGACTTCGCGCCAGGTCAGAACCCCAAGACAGATGGAGCAAACTGCATTG	300			
OY	101 AlAProPogInPrcaIalyvgInPhelenuIeserProbosserPrProvaIgIy	120			
Db	301 GCACCTCCACAGCGCTGCCAAACAGTTCTCATATTCACACCCTTCCTCCCTGTGGC	360			
OY	121 TrPlvysProIleserAspaIatnPrProValIeuAntyrAspleuLeutyralaVala	140			
Db	361 TGAAGGCTATTCAGCGATTCGCACACAGCTCTCAATATGACTTATTATGGCTGTGGC	420			
OY	141 LyvAleuGlYProglYgulytyrgJuleuHieaglYthrgrIsusrThrProserVal	160			

D	421	AAATAGGACGACGAGAGAATAATGACTCAGCGCTGGAACTGAGTGTACACCAGGCCTC	480
OY	161	VAlVAlHlSVAlCYsApPSeRAsPMeGlUGlUGlUAPPrOlySThrSerProLyS	180
D	481	GtGTGTGAtGTGTGTGACACGACATGAGAGAGAGAGACCCMAAGACTTCGCCCAAG	540
OY	181	ProlYSIleIllegITnTrArGArgProGIyLeuProPSeRValSerAn	197
D	541	CCAAAATCATTCAGACCCGGCGTCCGGGCTTGCCACCTCCGTGTCAAC	591
RESULT 2			
LOCUS	AF237791	594 bp	mRNA linear ROD 12-Apr-2000
DEFINITION	Mus musculus myocyte-enriched calcineurin interactin protein 2		
ACCESSION	AF237791		
VERSION	AF237791.1	GI:7542530	
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus.		
REFERENCE	Bukacynski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus		
AUTHORS	Rocheamel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.		
TITLE	A Protein Encoded Within the Down Syndrome Critical Region Is Enriched in Striated Muscles and Inhibits Calcineurin Signaling		
JOURNAL	J. Biol. Chem. (2000) In press		
REFERENCE	2. (bases 1 to 594)		
AUTHORS	Rocheamel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
FEATURES			
Source	1..594	/organism="Mus musculus"	
CDS	1..594	/db_xref="taxon:10090"	
	/note="MCIIP2; similar to Homo sapiens ZAK14; similar to EST sequence of Genbank Accession Number AU035927"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin interactin protein 2"		
	/protein_id="AAF63487.1"		
	/db_xref="GI:7542531"		
	/translation="MPASMDQVSTLVACVDVEFTNOBVKKEFGLFKTYDECVTL FOLKFSRRVRINFSHPKMSARALIEHQRFKKLKLPAQVGTEPDODDKHLAR POKPKDFLEPPSPPYGKKRIKSIAETPVLYNDLLNAAKGGRKELHNAGTESTPSV VHVCDSDMEEDBEEDPKTSVPKRKIQTNRPGULPVSNN"		
BASE COUNT	148 a	165 c	149 g 132 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.49e-84	Length:	594
Score:	1041.00	Matches:	197
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-782-953-9 (1-197) x AF237791 (1-594)			
OY	1	McCPoAlAProSeRmeCAsPCysAsPvAlSeThrLeuValAlAcyeValValAsPvAl	20
D	1	ATGCCAGCCCCCTAGCATGACCTGTATTGTTCCACTCTGCTGCTGTGTGATG	60
OY	21	GUVALPhetThAsGInGluValylSGluTyphetiGluGlyLeuPheArgrThrTyraP	40
D	61	GAGGCTTTTACCAATCAGAGAGGTTAAGAAAAATTGAGAGGACTGTCGGAACCTATGAT	120
OY	41	GUcYevAlThPhetiGlnleuPheHySerPheAdgArGvAlArgIleasnPheserHis	60

```

Db 121 GAATGTGTGAGCTTCCAGCTGTTTAAGAGTTTCCGACGGTTCGAATAAATTTTCAGCCAT 180
Qy 61 ProlyssSerAlaAlaAaArgileGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 181 CCCAAATCTGCAGCCCGTGCAGATAGAGCTTATGAGACTCAGTTCAGAGGGAAGAAG 240
Qy 81 LeuLysLeuTyrrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 241 CTAAAACTCTACTTTCGCCAGGTCCAGACCCAGAGACAGATGAGACAAACTGCAATTTG 300
Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 301 GCACCTCCAGAGCTGCCAAACAGTTCCTCATCTCACCCCTTCATCTCTCTCTCTCTCT 360
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrrAspLeuLeuTyrrAlaValAla 140
Db 361 TGAAGCCTATCAGCATGACACACCATGCTCACTCACTATGACCTTCTTATGCTGTGCC 420
Qy 141 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 421 AACTAGAGCAGGAGAGAAATATGAGCTGCACGCTGGAACCTGAGTCTACACCGAGGTC 480
Qy 161 ValValHisValCysAspSerAspMetGluGluGluGluAspProTyrrSerProLys 180
Db 481 GTGGTGCATGTGTGACGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 541 CCAAAATCATTCAGACCCCGCGTCCGGGCTTGCCACCCCTCCGTTCCAC 591

RESULT 3
AF237887 659 bp mRNA linear ROD 17-NOV-2000
LOCUS Mus musculus Down syndrome candidate region 1-like protein 1
DEFINITION (Decr11) mRNA, complete cds.
ACCESSION AF237887
VERSION AF237887.1 GI:7417330
KEYWORDS Mus musculus.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 659)
Tiripoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
TITLE The murine DSCR1-like (Down syndrome candidate region 1) gene
JOURNAL family: conserved synteny with the human orthologous genes
MEDLINE Gene 257 (2), 223-232 (2000)
PUBMED 11080588
REFERENCE 2 (bases 1 to 659)
Tiripoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
AUTHORS Direct Submission
TITLE Submitted (22-FEB-2000) Istituto di Istologia ed Embriologia
JOURNAL Generale, Università di Bologna, Via Belmeloro, 8, Bologna, BO
40126, Italy
FEATURES
Location/Qualifiers
1..659
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="17"
/sex="male"
/tissue_type="brain"
1..659
/gene="Decr11"
47..640
/gene="Decr11"
/note="Decr11"
region 1-like protein 1"
/codon_start=1
/product="Down syndrome candidate region 1-like protein 1"
/protein_id="AAF62538.1"

```

```

/db_xref="GI:7417331"
/translation="MPAPSMDCDVSTLVACVVDVFTNQEVKEPEGLFRYDECVT
FQFMSFRVRINFPKSAARIELHETQFRGKGLYFAQVPTPTDGGKHLAP
POPAKOFILSPSPVGVKPISTADPLVNTDLYAVAKLGPBKRYELHAGTSPSV
VHVDCSDMEEDPKTPKPIITRRPGLPPSVN"
BASE COUNT 167 a 183 c 163 g 146 t
ORIGIN
Alignment Scores: 2.79e-84 Length: 659
Score: 1041.00 Matches: 197
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-782-953-9 (1-197) x AF237887 (1-659)
Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 47 ATGCCAGCCCCCTAGCATGGACTGTGATGTTTCCACTCTGCTGCTGCTGCTGCTGCTGCTG 106
Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrrAsp 40
Db 107 GAGGTCTTTTACCAATCAGGAGGTAAAGGAAAAATTCAGGGGACTGTTCGGGACCTATGAT 166
Qy 41 GluCysValThrPheClnLeuPheLysSerPheArgValArgValArgLysPheSerHis 60
Db 167 GAATGTGTGACCTTCCAGCTGTTTAAAGAGTTTCCAGCGGTTTCSAATAAATTTACGCCAT 226
Qy 61 ProLysSerAlaAlaArgAlaArgileGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 227 CCCAAATCTGCAGCCCGTCCCGGATAGAGCTTCATGAGACTCAGTTCAGAGGGAAGAAG 286
Qy 81 LeuLysLeuTyrrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 287 CTAAAACTCTACTTCCGCCAGGTCCAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346
Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 347 GCACCTCCAGAGCTTCCCAACAGTTCCTCATCTCACCCCTTCATCTCTCTCTCTCTCT 406
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrrAspLeuLeuTyrrAlaValAla 140
Db 407 TGAAGCCTATCAGCGATGTCACACCATGCTCTCACTATGACCTTCTTTATGCTGTGGCC 466
Qy 141 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 467 AAACCTAGGAGCAGGAGAGAAATATGAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAG 526
Qy 161 ValValHisValCysAspSerAspMetGluGluGluGluAspProLysThrSerProLys 180
Db 527 GTGGTGCATGTGTGACAGCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 586
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 587 CCAAAAATCATCCAGACCCCGCTCCCGGCTGCCACCCCTCCGTTCCCAAC 637

RESULT 4
AB061524 3227 bp mRNA linear ROD 13-NOV-2001
LOCUS Mus musculus mRNA for calcineurin inhibitory protein ZAKI-4,
DEFINITION complete cds.
ACCESSION AB061524
VERSION AB061524.1 GI:16904644
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Kanou, Y., Miyazaki, T., Seo, H. and Murata, Y.
TITLE calcineurin inhibitory protein ZAKI-4

```



Db 670 \*GTCGTGCACGGTGTGGCAGCTGATCATAGAGAAGAGGACCACCAAGACTTCCCAAAG 729

Qy 181 ProLyslelleGlnThrArgProGlyLeuProSerValSerAsn 197  
|||||  
Db 730 CCAAAAATCATCCAATACTCGCGCTCTGGCTGCCACCCTCCGTGTCCAAC 780  
|||

RESULT 6  
AX333015                      3184 bp    DNA       linear    PAT 09-JAN-2002

LOCUS  
DEFINITION Sequence 3524 from Patent WO0194629.  
ACCESSION AX333015  
VERSION AX333015.1 GI:18123649  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets

JOURNAL Patent: WO 0194629-A 3524 13-DEC-2001;  
Avalon Pharmaceuticals (US)

FEATURES  
source Location/Qualifiers  
1..3184 /organism="Homo sapiens"  
BASE COUNT 921 a -681 c 657 g 925 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.7e-80 Length: 3184  
Score: 1007.00 Matches: 190  
Percent Similarity: 98.98% Conservative: 5  
Best Local Similarity: 96.45% Mismatches: 2  
Query Match: 96.73% Indels: 0  
DB: 6 Gaps: 0

US-09-782-953-9 (1-197) x AX333015 (1-3184)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValaspVal 20  
|||

Db 190 ATGCAGCCCCATCATGTGACGTGTGAATTTCACCTTGTTGCTGTGTGGTAGATGC 249  
|||

Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrrAsp 40  
|||

Db 250 GAGGCTCTTTACCAATCAGGAGTTTAAGAAAAATTTGGGGGACTGTTTCGGACTTATGAT 309  
|||

Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgIleAsnPheSerHis 60  
::: |||

Db 310 GACTGTGTGAGCTTCCAGCTATTTAAAGAGTTTCAGAGCTGTCCGTATATAAACCTTCAGCAAT 369  
|||

Qy 61 ProllysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80  
|||

Db 370 CCTAAATCTGAGCCCAGCTAGATAGAGCTTCATGAACCCCAATTCAGAGGGAAAAA 429  
|||

Qy 81 LeuLysLeuTyrrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
|||

Db 430 TTTAAAGCTCTACTTTGCACAGTTTCAGACTCCAGACACAGATGGAGACAACTGCACCTTG 489  
|||

Qy 101 AlaProGlnProAlaLysGlnPheLeuIleSerProProSerSerProValGly 120  
|||

Db 490 GCTCACCCCCAGCTGCCAAACAGTTTTCTATCTTCGGCCCCCTTCCTCCACCCTGTTAGC 549  
|||

Qy 121 TrpLysProfileSerAspAlaThrProValLeuAsnTyrrAspLeuLeuTyrrAlaValAla 140  
::: |||

Db 550 TGGCAGCCCCATCAGCATGCCACGCCAGTCTCTCAACTATGACCTCTCTATGCTGTGGCC 609  
|||

Qy 141 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaGlyThrGluSerThrProSerVal 160  
|||

Db 610 AAATAGGACCAGGAGAGAGTATGAGCTTCATCCAGGACTGATGTCACCCCAAGTSTC 669  
|||

	Qy	670	GTCGTGCACGTTGGCAGCACTGATAGAGGAAGAAGACCACCCAAAGACTTCCCCAAG	729
	Db	181	ProlysllellleGlThrArgArqProGiLyLeuProSerValSerAsn	197
	Db	730	CCAAATAATCATCCAACATCGCGCTTGCTGCCCTGCCACCTCCGTGTCCAAC	780
RESULT 7				
AX365327				
LOCUS		AX365327	3184 bp DNA linear PAT 15-FEB-2002	
DEFINITION		Sequence 17 from Patent WO0204491.		
ACCESSION		AX365327		
VERSION		AX365327.1 GI:18697053		
KEYWORDS				
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		Williams,S.R. and Rothermel,B. Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)		
JOURNAL		Patent: WO 0204491-A 17 JAN-2002; Board of Regents, The University of Texas System (US); Williams, Sanders R. (US) ; Rothermel, Beverly (US)		
FEATURES		Location/Qualifiers		
source		1..3184		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
CDS		205..783		
		/notes="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAD23813.1"		
		/db_xref="GI:18697054"		
		/translations="MDCDVSTLVCVVDVEFTNQBVKEKPGGLFRYYDDCVTFQLFKP SPRRVRINFSNPKSAARAIELHFOTGRKKLKLYFAOQTPTDGDGKLHLAPPOPAK QFLISPPSPPVSWQPINDPATPNYLDLLAVAKLGEGEKVELHAGTESTPSVVHVHC DSIDIEEDDPKTSPKPILIIQTRRPLPPSVSN"		
BASE COUNT		921 a 581 c 657 g 925 t		
ORIGIN				
Alignment Scores:				
Pred. No.:		Length: 3184		
Score:		Matches: 190		
Percent Similarity:		Conservative: 5		
Best Local Similarity:		Mismatches: 2		
Query Match:		Indels: 0		
DB:		Gaps: 0		
US-09-782-953-9 (1-197) x AX365327 (1-3184)				
Qy	1	MetProAlaProSerMetAspValSerThrLeuValAlaCyseValValaspVal	20	
Db	190	ATGCACGCCCTTAGCATGGACTGTGATGTTTTCCACTCTGGTTGGCTGTGGTGCATGTC	249	
Qy	21	GluValPheThrAnGlnGluValIlysGluLylsPheGluClgLyLeuPheArqThrTyASP	40	
Db	250	GAGTCTTTTACCATCACGAGGTTAAGAAAATTTCGGGGACGTGTTTCGACATTATGAT	309	
Qy	41	GluCyseValThrPheGInLeuPheLyseserPheArgArgValArgIlleaenPheserHis	60	
Db	310	GACTGTGTGAGCTTCCAGCTATTTAAGAGYTTCAGACGTGTCGTATAACTTCAGCAAT	369	
Qy	61	ProtlyserAlalaArgAlaArgIlleGluLeuHisGluThrGlnPheArqGilylslays	80	
Db	370	CCTAAATCTCAGCCCCAGCTAGGATAGCTTCATGAAACCCCATTTCAGAGGGAAAAA	429	
Qy	81	LeulystleutyrrPheaIGlnValGlnThrProGluThràspGlyAspylsLeuHiLeu	100	
Db	430	TAAAGTCTACTTTGCACAGGTTTCAGCTCCAGACACATGGAGACAACACTGCATTG	489	
Qy	101	AlaProGlnProAlaLylsGlnPheLeuIlleserProProSerSerProProValgly	120	







RESULT 11  
AX420427/c 1021 bp DNA linear PAT 18-JUN-2002  
LOCUS  
DEFINITION Sequence 3 from Patent WO0204513.  
ACCESSION AX420427  
VERSION AX420427.1 GI:21524582  
KEYWORDS  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1  
AUTHORS Loring, J.F., Tingley, D.W., Edwards, C.M. and Screeter, D.G.  
TITLE Down syndrome critical region 1-like 1 proteins  
JOURNAL Patent: WO 0204513-A 3 17-JAN-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1. 1021  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/note="Incyte ID No: 219435\_Rn.1"  
BASE COUNT 237 a 241 c 286 g 257 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,266-72 Length: 1021  
Score: 910.00 Matches: 192  
Percent Similarity: 94.63% Conservative: 2  
Best Local Similarity: 93.66% Mismatches: 3  
Query Match: 87.42% Indels: 8  
DB: Gaps: 0  
US-09-782-953-9 (1-197) x AX420427 (1-1021)  
QY 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20  
DB 855 ATGCCAGCCCTAGCATGATGATGTTCCACTCTGTCGCTGTGTGTGATGG 796  
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyraa 40  
DB 795 GAGGCTTTTACCAATCAGAGGTTTAAAGAAATTTGAGAGAGCTGTTCCGACCTATGAC 736  
QY 41 GluCyValThrPheGlnLeuPheLysSerPheArgArgValArgIleAsnPheSerH 60  
DB 735 GATGTGTGACCTTTACGCTGTTTAAAGCTTCCAGCGGTTCCGAAATTAATTCAGCA 676  
QY 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHISGlu-ThrGlnPheArg-GlyL 79  
DB 675 CCCCAAGCTGCAGCCCGGTGCCGATAGGAGCTTCATGTGAGCCACCTCATGAGGGA 616  
QY 79 ValValLeuLysLeuLysPheAlaGlnValGlnThrProGlnThrAspGly-AspLysLeu 98  
DB 615 ACAAAGCTGAACCTTACTTTGACACAGTCCAGACCCCAAGACAGATGATGACCAACTG 556  
QY 99 His-LeuAlaProProGlnProAlaLysGlnPheLeuLieserProPheSerSerProBr 118  
DB 555 CAGTTTGGACACCCCAACACCTGCGCAACATTCCTCATCTCACCCCTTCACTCTCC 496  
QY 118 oValGlyTpyLysProLieserAspAlaThrProValLeuAsn-TyraaPheLeuTyra 138  
DB 495 CATTGGCTGGAAGCCTATCAGCGATCCACACCAAGTCTCAAGTACGACCTCTTATG 436  
QY 138 ValValAlaLysLeuGlyProGlyGlyLysTyrgLLeuHISAlaGlyThrGlnLysThrP 158  
DB 435 CCCTGGCCCAACTAGACAGCAGAGAAATATGAGCTGATGCGGGAAGTGAAGTACAC 376  
QY 158 roSerValValValHisValCyAspSerAspMetGluGluGluGluAspProLysThs 178  
DB 375 CAGAGCTTGTGTCACGCTGTGTGACAGCACTTGAAGAGAGAGAGATCCAAACACTT 316  
QY 178 sProLysProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197

DB 315 CCCCAAGCCAAATATCATCCAGACCCGCGCTCTGCGCTGCCACCTCGCTTCAC 257  
RESULT 12  
AX074351 934 bp DNA linear PAT 06-FEB-2001  
LOCUS  
DEFINITION Sequence 3 from Patent WO0104305.  
ACCESSION AX074351  
VERSION AX074351.1 GI:12710514  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1  
AUTHORS Tang, Y.T. and Yue, H.  
TITLE Human proteins involved in detoxification  
JOURNAL Patent: WO 0104305-A 3 18-JAN-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1. 934  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 247 a 244 c 206 g 237 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,73e-72 Length: 934  
Score: 905.00 Matches: 171  
Percent Similarity: 95.72% Conservative: 8  
Best Local Similarity: 91.44% Mismatches: 8  
Query Match: 86.94% Indels: 0  
DB: Gaps: 0  
US-09-782-953-9 (1-197) x AX074351 (1-934)  
QY 11 SerThrLeuValAlaCyValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
DB 161 AACTCGTGTGTTGCTGTGAATGATTCACACGTCAGTGTGTAAGAGAGAGACCAAGAA 220  
QY 31 LysPheGlnGlyLeuPheArgThrTyraaPheGluCyValThrPheGlnLeuPheLysSer 50  
DB 221 AAATTTGAGGACACTGTTCCGACCTTATGATACCTGTGTGAGCTTCACCTATTAAAGCT 280  
QY 51 PheArgArgValArgIleAsnPheSerHisAspProLysSerAlaAlaArgAlaArgIleGlu 70  
DB 281 TTCACAGCTGTCCGATATTAACCTTCAGCAATCTTAATCTGCAGCCGAGCTAGATACAG 340  
QY 71 LeuHISGluThrGlnPheArgGlyLysValLeuLysLeuLysPheAlaGlnValGlnThr 90  
DB 341 CTTCAATGAAACCCCAATTCAGAGGGAATAATTAAGCTCTACTTTCACAGGTTCCAGACT 400  
QY 91 ProGlnThrAspGlyAspLysLeuHISLeuAlaProProGlnProAlaLysGlnPheLeu 110  
DB 401 CCAGAGACAGATGAGACAAACTGCACTTGCTCCACCCAGCTGCCAAACAGTTTCTC 460  
QY 111 LieserProProSerSerProProValGlyTpyLysProLieserAspAlaThrProVal 130  
DB 461 ATCTGCCCTCTCTCTCCCACTGTTGGCTGGCGAGCCCAACAGATGACCGCCAGTGC 520  
QY 131 LeuAsnTyraaPheLeuLysTyraValAlaLysLeuGlyProGlyGlyLysTyrgLLeu 150  
DB 521 CTCAACTATGACCTCTCTATGCTGTGCGCAAACTAGAGACAGAGAAAGTATAGCTTC 580  
QY 151 HisAlaGlyThrGlnLysThrProSerValValValHisValCyAspSerAspMetGlu 170  
DB 581 CATGAGGAGACTGATCCACCCCAAGTGTGTCGACGCTGTGACAGTGAACATGAGG 640  
QY 171 GluGluGluAspProLysThsSerProLysProLysIleIleGlnThrArgArgProGly 190  
DB 641 GAAAGAGAGAGACCAAACTTCCCAAGCCAAATAATCATCAAACTCGGCTCTGCTG 700  
QY 191 LeuProProSerValSerAsn 197



Db 701\*CTGCCACCTCCGTGTCCAAC 721

RESULT 13  
AX420425  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT  
ORIGIN

AX420425  
Sequence 1 from Patent WO204513.  
AX420425  
AX420425.1 GI:21524581  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Loring, J.F., Tingley, D.W., Edwards, C.M. and Streeter, D.G.  
Down syndrome critical region 1-like 1 proteins  
Patent: WO 0204513-A 1 17-JAN-2002;  
Incyte Genomics, Inc. (US)  
Location/Qualifiers  
1..3159  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/notes="Incyte ID No: 247500.5"  
926 a 657 c 657 g 915 t 4 others

Alignment Scores:  
Pred. No.: 2,17e-71 Length: 3159  
Score: 905.00 Matches: 171  
Percent Similarity: 95.72% Conservative: 8  
Best Local Similarity: 91.44% Mismatches: 8  
Query Match: 86.94% Indels: 0  
DB: 6 Gaps: 0

US-09-782-953-9 (1-197) x AX420425 (1-3159)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
Db 389 AACTCGTGTTCGGTCAATGTTCCAGTCAGTGTGACGTTCCAGCTATTTAAGAGT 508  
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
Db 449 AAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGT 508  
QY 51 PheArgValArgLeuAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70  
Db 509 TTCAGAGGTGTCGGTATAAACTTCAGCAATCTTAATCTCGACGCCGAGCTAGGATAGAG 568  
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
Db 569 CTTATGAAACCCCAATTCAGAGGAAAAATTAAGCTCTTACTTTGCACAGGTTTCAGACT 628  
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110  
Db 629 CCAGACAGATGAGACAACTGCATTCGCTCCACCCAGAGCTGCCAAACAGTTTCTC 688  
QY 111 IleSerProSerSerProProValGlyTyrPheProLysSerAspAlaThrProVal 130  
Db 689 ATCTGCCCCCTTCCTCCCCACCTGTTGGTGGCAGCCCATCAACGATGCCAGCCAGTC 748  
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
Db 749 CTCACTATGACCTCTCTATGCTGTGGCCAACTAGGACGAGAGAGATATAGACTC 808  
QY 151 HisAlaGlyThrGluSerThrProSerValValValHisValCysAspSerAspMetGlu 170  
Db 809 CATGCAGGAGCTGAGTCCACCCCAAGTGTGTCGACGCTGCCACAGTGCATAGAG 868  
QY 171 GlyGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190  
Db 869 GAAGAAGAGAGACCAAGACTTCCCAAGCCAAAAATCATCAAACTCGGCGCTCTGGC 928  
QY 191 LeuProProSerValSerAsn 197

Db 929 CTGCCACCTCCGTGTCCAAC 949

RESULT 14  
AK090990  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
BASE COUNT  
ORIGIN

AK090990  
Homo sapiens cDNA FLJ33671 fis, clone BRAWH2001459, highly similar to ZAKI-4 PROTEIN.  
AK090990  
AK090990.1 GI:21749260  
oligo capping; fis (full insert sequence).  
Homo sapiens brain cDNA to mRNA, clone\_lib:BRAWH2  
clone:BRAWH2001459.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuo, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 3261)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.  
Location/Qualifiers  
1..3261  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="BRAWH2001459"  
/tissue type="brain"  
/clone\_lib="BRAWH2"  
/note="cloning vector: pME18SFL3"  
943 a 710 c 679 g 929 t

Alignment Scores:  
Pred. No.: 2,24e-71 Length: 3261  
Score: 905.00 Matches: 171  
Percent Similarity: 95.72% Conservative: 8  
Best Local Similarity: 91.44% Mismatches: 8  
Query Match: 86.94% Indels: 0  
DB: 9 Gaps: 0

US-09-782-953-9 (1-197) x AK090990 (1-3261)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
Db 293 AACTCGTGTTCGGTCAATGTTCCAGTCAGTGTTCAGGAGAGAGAGCAGGAGNA 352  
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
Db 353 AAAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTCAGCTTCCAGCTATTTAAGAGT 412  
QY 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70  
Db 413 TTCAGACGTGTCGGTATTAACCTTCAGCAATCTTAATCTCGAGCCCGAGCTAGGATAGAG 472

QY	71		LeuHsGluThrGlnPheAspIGLYLyleValLeuValLeuTYrPheAlaGlnValGlnThr	90
Db	473		CTTCATGAAGAACCCCAATTGCAGAGGAAAAAATTAAAGCTCTACTTTGGACAGGTTCACACT	532
QY	91		ProGUThrAspGlyAspLyLeuHisIleuLaIProFgInProAlaTyGInPheLeu	110
Db	533		CCAAGACAGATGGAGCAACAATCGACTTGGCTCACCACCGAGCTGCCCAACACTTCTC	592
QY	111	IleSerProFroSerSerProProValGlyTYrplySerProlIleSerAspAlaThrProVal	130	
Db	593	ATCTGCCCCCTTCTCCCTCCACCTGTGGCTGGGAGCCCAACAGATGCCACCGCAGTC	652	
QY	131	LeuValTYrAspLeuLeuTYrAlaValAlaTyLeuGlyPYrProGlyGluTYrGlyIleu	150	
Db	653	CTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGAGAGAGAGATATGAGCTC	712	
QY	151	H1sAlaGlyThyGlnSerThyProSerValValHisValCYsAspSerAspMetGlu	170	
Db	713	CATGAGGAGCTAGATGCCACCCCAAGTGTCGTGGAGCGTGGCAGATGACATGAGAG	772	
QY	171	GluGluGluAspProGlyThySerProLySerProLyIleIleGlnThrArgArgProGly	190	
Db	773	GAGAAGAGGAGCACCAAACTCCCACAAAGCCAAAATAATCATCCAACTCGGCGTCTGGC	832	
QY	191	LeuProProSerValSerAsn	197	
Db	833	CTGCCACCCCTCGTGTCCAAC	853	
RESULT	15			
LOCUS	AY034085	3240 bp	mRNA	linear PRI 11-JUL-2002
DEFINITION			Homo sapiens calcineurin inhibitor ZAKI-4 beta splice variant 1	
ACCESSION	AY034085		mRNA, complete cds, alternatively spliced.	
VERSION	AY034085.1	GI:21307622		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 3240)			
AUTHORS	Gao,X., Kambe,F., Miyazaki,T., Sarkar,D., Ohmori,S. and Seo,H.			
TITLE	Novel human ZAKI-4 isoforms: Hormonal and tissue-specific regulation and function as calcineurin inhibitors			
JOURNAL	Biochem. J. (2002)			
PUBMED	12102656			
REFERENCE	2 (bases 1 to 3240)			
AUTHORS	Gao,X., Kambe,F., Miyazaki,T., Ohmori,S. and Seo,H.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-MAY-2001) Endocrinology and Metabolism, Research Institute of Environmental Medicine, Nagoya University, Furo-cho, Nagoya 464-8601, Japan			
FEATURES	Location/Qualifiers			
source	1..3240			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="6"			
	/map="gp12"			
	/tissue_type="brain"			
	108..839			
	/note="alternatively spliced"			
	/codon_start=1			
	/product="calcineurin inhibitor ZAKI-4 beta splice variant 1"			
	/protein_id="AAK59805.1"			
	/db_xref="GI:21307623"			
	/translation="MRGSYSYTRGRSSPGCGHVHDGSLFLCCIDRDMVTRPARAEAFQATDFNDLPNSLFACNVHQSVFEGBSEKEKGGLFRYYDDCVYTQLKSPRRVITNFSNPASARAARIELHETPRFGKKLKIFYFAVGOTPETDGDHLAPQPAPKPLISPPSPSPSWOPINDATPVNLVDLVAVAKLGPEKELHAGTESPTSVVVHVCDSIDIEEDPKSPKPKIIQRRLPLPSVSN"			
	333..3340			
	/note="similar to sequence with GenBank Accession Number misc_feature			

	polyA_signal	D83407"	1079. .1084
	polyA_signal	/note="alternative"	
BASE COUNT	947 a 688 c 677 g 928 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.23e-70	Length:	3240
Score:	892.00	Matches:	169
Percent Similarity:	94.65%	Conservative:	8
Best Local Similarity:	90.37%	Mismatches:	10
Query Match:	85.69%	Indels:	0
DB:	9	Gaps:	0
US-09-782-953-9 (1-197) x AY034085 (1-3240)			
Qy	11 SerthrlenuValAlaCyveValAlaAspValGluValPheThrAsnGlnGluValLysGlu	30	
Db	276 AACCTGCTGTTGGCTGATGTCACACGCTCACTGTTTAAAGAGAGAGAGACCAAGAA	335	
Qy	31 LysPheGlnGlyLeuPheArgThrTyAspGluCyveValThrPheGlnLeuPheLysSer	50	
Db	336 AAATTGGGGGACGTTTGGACTTATGATGACTGTGTGACGTTCCAGCTATTAAAGAGT	395	
Qy	51 PheArgValAlaGlyLeuPheSerHisProLysSerLalaAlaArgAlaGlyLeu	70	
Db	396 TTCAGACGTGTCGGATTAACTTACGAATCTTAATCTCAGCCGAGCTAGATAGAG	455	
Qy	71 LeuHisGlnThrGlnPheArgGlyLysLysLeuLysLeuTyPheAlaGlnValGlnThr	90	
Db	456 CTTCAATGAACCCCAATTCAGAGGAAAAATTAAAGCTTCACTTGACAGGTTCAAGACT	515	
Qy	91 ProGlnThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu	110	
Db	516 CCAGAGACAGATGAGAGCAAACTCACTGGCTCCACCCAGCCTGCGCAACAGTTTCTC	575	
Qy	111 IleSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVal	130	
Db	576 ATCTGCCCCCTCTCTCCCACTCTTAGCTGAGAGCCCAATCAAGATGCCACCCAGTGC	635	
Qy	131 LeuAsnTyraaPLeuLeuTyraAlaValAlaLysLeuGlyProGlyGlyLysTyGlyLeu	150	
Db	636 CTCACACTATGACCTCTCTTAGCTGCTGGCCAAACTAGACACGAGAGAGAGTATGACCTC	695	
Qy	151 HisAlaGlyThrGlnGlySerThrProSerValValAlaHisValCysAspSerAspMetGlu	170	
Db	696 CATGGAGGAGACTGATGCCACCCCAAGTGTCTCTGACAGGTGCGACAGTGACATAGAG	755	
Qy	171 GlnGlnGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly	190	
Db	756 GAAGAAAGAGAGACCCAAAGACTTCCCAAGAGCCAAATAATCATCAAACTGGCGTCTGGC	815	
Qy	191 LeuProProSerValSerAsn	197	
Db	816 CTGCCACCTCCGTGTCCAC	836	

Search completed: December 14, 2002, 21:32:11  
Job time : 1654.42 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1112.9 Seconds  
(without alignments)  
2866.843 Million cell updates/sec

Title: US-09-782-953-9  
Perfect score: 1041  
Sequence: 1 MFAPMDCDVSTLVACVDV.....SPKPKIIQRRRGLPPSVSN 197

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2\_1/USPTO spool/US09782953/runat\_11122002\_114429\_17353/app query.fasta\_1.1173  
-DB=EST -OFMT=fasta -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09782953 @CNC 1.1583 @runat\_11122002\_114429\_17353 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gas: \*  
18: em\_gas\_hum: \*  
19: em\_gas\_inv: \*  
20: em\_gas\_pln: \*  
21: em\_gas\_vrt: \*  
22: em\_gas\_fun: \*  
23: em\_gas\_mam: \*  
24: em\_gas\_mus: \*  
25: em\_gas\_other: \*  
26: em\_gas\_pro: \*  
27: em\_gas\_rod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	98.0	925	14	BQ720798
2	1007	96.7	1055	14	BM925558 AGENCOURT
3	994	95.5	903	9	AL532776
4	941.5	90.4	847	13	BI545609
5	854	82.0	753	10	AW957141
6	849	81.6	758	12	BG709000
7	805	77.3	495	12	BF443461
8	803	77.1	742	12	BG721994
9	788	75.7	546	14	BQ260356
10	786.5	75.6	617	13	BI476888
11	774	74.4	640	10	BB650331
12	758.5	72.9	1007	12	BF534085
13	748	71.9	557	9	AI796134
14	728	69.9	626	12	BG701025
15	726	69.7	747	9	AI69846
16	715	68.7	516	9	AL693037
17	674.5	64.8	1103	11	AK019377
18	669	64.3	530	9	AA814235
19	658.5	63.3	863	14	BQ438012
20	655.5	62.8	829	9	AU035927
21	646	62.1	550	12	BG80881
22	641	61.6	837	9	AU124628
23	641	61.6	931	9	AL543576
24	641	61.6	939	9	AL546617
25	641	61.6	1041	13	BM450020
26	636	61.1	946	14	BQ278576
27	635	60.9	718	9	AU131040
28	634	60.9	662	13	BI464521
29	630.5	60.6	1158	14	BQ895506
30	630	60.5	642	14	BQ520533
31	630	60.5	689	9	AL652662
32	630	60.5	748	14	BM964376
33	628	60.3	433	9	AA248590
34	628	60.3	885	9	AL544755
35	626.5	60.2	828	9	AL536447
36	626.5	60.2	867	14	BQ427531
37	626.5	60.2	890	9	AL538796
38	626	60.1	1078	13	BM541636
39	615	59.1	939	9	AL554686
40	614	59.0	898	14	BQ733563
41	613.5	58.9	863	14	BQ894646
42	612	58.8	906	9	AL544313
43	606	58.2	599	13	BQ064020
44	604	58.0	664	13	BJ061844
45	601	57.7	596	13	BJ058360

# ALIGNMENTS

RESULT 1  
BQ720798  
LOCUS  
DEFINITION AGENCOURT\_8230839 Lupski\_sympathetic\_trunk Homo sapiens cDNA clone  
IMAGE:619T156 5', mRNA sequence.  
ACCESSION BQ720798  
VERSION BQ720798.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 925)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40  
 Db 252 GAGGTCCTTTACCAATCAGAGGTTAAGAAATTTAGGAGACTGTTTCGGACTTATGAT 311

Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgLysAsnPheSerHis 60  
 Db 312 GACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGTCGGTATAACTTCAGCAAT 371

Qy 61 ProllysSerAlaAlaArgAlaArgLysGluLysGluLysGluLysGluLysGluLys 80  
 Db 372 CCTAAATCTGCAGCCGAGCTAGGATAGCTTCATGAACCCCAATTCAGAGGGGAAAAA 431

Qy 81 LeuLysLeuTyPheAlaGlnValGlnThrProGlnThrAspGlyAspLysLeuHisLeu 100  
 Db 432 TTAAGCTCTACTTTGACAGGTTTCAGACTCCAGAGACAGATGAGACAACTGCACCTG 491

Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120  
 Db 492 GCTCAGCCAGCTGCCAACAGTTTCTCATCTGGCCCTTCTCTCCACCTGTTGGC 551

Qy 121 TrpLysProlLysSerAspAlaThrProValLeuAsnTyRAspLeuLeuTyRAlaValAla 140  
 Db 552 TGGCAGCCCATCAACGATGCCACGCTCTCACTATGACCTCTCTATGCTGTGGCC 611

Qy 141 LysLeuGlyProGlyGluLysTyRLeuHisAlaGlyThrGluSerThrProSerVal 160  
 Db 612 AACTAGACACGAGAGAGATGAGCTCCATGACGGACTGAGTCCACCCCAAGTGC 671

Qy 161 ValValHisValCysAspSerAspMetGluGluGluGluAspProLysTyRSerProLys 180  
 Db 672 GTCTGCAGCTGTCGACAGTGACATAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 731

Qy 181 ProlLysLeuGlnThrArgArgProGlyLeu-ProProSerValSerAsn 197  
 Db 732 CCAAAATATCATCCAACTCGCGCTGCTGGCTGCCACCTCTCGTGCCAAAC 783

RESULT 3  
 AL532776

LOCUS  
 AL532776 LTI\_FL015\_Brnl Homo sapiens cDNA clone CS0DN001Y008 5  
 Prime, mRNA sequence.

ACCESSION  
 AL532776

VERSION  
 AL532776.1 GI:12796269

KEYWORDS  
 EST.

SOURCE  
 human.

ORGANISM  
 Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE  
 Full-length cDNA libraries and normalization

JOURNAL  
 Unpublished (2001)

COMMENT  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
 Location/Qualifiers

Source  
 i. .903  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DN001Y008"  
 /clone\_lib="LTI\_FL015\_Brnl"  
 /sex="male"  
 /tissue\_type="Adult brain"  
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with NotI and  
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6  
 vector. Library was constructed by Life Technologies.  
 Contact : Feng Liang Life Technologies, a division of  
 Invitrogen 9800 Medical Center Drive Rockville, Maryland  
 20850, USA Fax : (1) 301 610 8371 Email :  
 fliang@lifetech.com URL :

BASE COUNT 222 a 249 c 206 g 221 t 5 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.65e-98 Length: 903  
 Score: 994.00 Matches: 189  
 Percent Similarity: 98.48% Conservative: 6  
 Best Local Similarity: 95.45% Mismatches: 2  
 Query Match: 95.45% Indels: 1  
 DB: 9 Gaps: 0

US-09-782-953-9 (1-197) x AL532776 (1-903)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
 Db 171 ATGCCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGGTGGCTGTGTGGTGGATGC 230

Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40  
 Db 231 GAGGTCTTTACCAATCAGGAGGTTAAGGAAATTTGAGGGACTGTTTCGGACTTATGAT 290

Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgLysAsnPheSerHis 60  
 Db 291 GACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGTCGTATAAATTCAGCAAT 350

Qy 61 ProlLysSerAlaAlaArgAlaArgLysGluLysGluLysGluLysGluLysGluLys 80  
 Db 351 CCTAAATCTGCAGCCGAGACTAAGATAGAGTTCATGAACCCCAATTCARAGGGGAAAAA 410

Qy 81 LeuLysLeuTyRAspAlaThrProValLeuAsnTyRAspLysLeuHisLeu 100  
 Db 411 TTAAGCTCTACTTTGACAGGTTTCAGACTCCAGAGACAGATGAGAGACAACTGCACCTG 470

Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120  
 Db 471 GCTCCACCCAGACCTGCCAACAGTTTCTCATCTGGCCCTTCTCTCCACCTGTTGG 530

Qy 120 YTrpLysProlLysSerAspAlaThrProValLeuAsnTyRAspLysLeuTyRAlaValAl 140  
 Db 531 CTGGCAGCCCATCAACGATGCCACCCAGTCTCACTATGACCTCTCTATGCTGTGGC 590

Qy 140 aLysLeuGlyProGlyGluLysTyRLeuHisAlaGlyThrGluSerThrProSerVa 160  
 Db 591 CAACTAGGACCCAGGAGAGATGAGTCTCATGAGGAGCTGAGTCCACCCCAAGTGT 650

Qy 160 lValValHisValCysAspSerAspMetGluGluGluGluAspProLysTyRSerProLy 180  
 Db 651 CTTCTGTGACGCTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710

Qy 180 sProlLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 711 GCCAAATATCATCCAACTCGCGTCTGCTGGCTGCCACCTCTCGGTGTCCAAC 762

RESULT 4

BI545609

LOCUS

DEFINITION

603187809P1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5259147 5',

BI545609

ACCESSION

BI545609.1 GI:15432921

VERSION

EST.

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 847)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

847 bp mRNA linear EST 05-SEP-2001







/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."  
BASE COUNT 134 a 145 c 122 g 94 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,766-78 Length: 495  
Score: 805.00 Matches: 158  
Percent Similarity: 98.79% Conservative: 5  
Best Local Similarity: 95.76% Mismatches: 1  
Query Match: 77.33% Indels: 2  
Gaps: 0

US-09-782-953-9 (1-197) x BF443461 (1-495)

QY 20 ValGluValPheThrNngInGluValLyGluLeuPheGluGlyLeuPheArgThrTyr 39  
DB 2 GTGGAGGCTTTACCAACGAGGAGGTTAAGA-MAATTGAGGGCTATTCCGAGATTAT 60  
QY 40 AAGGluCyValThrPheGlnLeuPheLeuSerPheArgArgValArgIleAsnPheSer 59  
DB 61 GACGAGTGTGTGACGTTCCAGCTGTTTAAAGCTTCAGACGTCTCCGATTAACTTCAGC 120  
QY 60 HsProlySerAlaAlaArgAlaArgIleGluLeuHsGluThrGlnPheArg-GlyLy 79  
DB 121 AATCCCAATCTGCACCCGAGCCAGATAGAGCTCATGMAACCCAGTTCAAGGGGGA 180  
QY 79 GlyLeuLeuLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLeuHs 99  
DB 181 AAAATTAAACTTACTTATGACAGGTTCCAGACTCCAGACAGACGAGGAGCAACTGCA 240  
QY 99 GLeuAlaProGlnProAlaLeuGlnPheLeuIleSerProPheSerProProVa 119  
DB 241 TTGGGCCCAACAGAGCTCCCAACAGTTCCATCTCACCCCGCTCTCTCCGT 300  
QY 119 GgTTPLyProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaVa 139  
DB 301 CGGCTGCAGCCCATCAGGATCCAGCGCGTCTTAACTAAGACCTCTATGCCGT 360  
QY 139 LAlaLeuLeuGlyProGlyGlyLeuTyrGluLeuHsAlaGlyThrGlnSerThrProSe 159  
DB 361 GCCCAAACTAGAGCAGAGAGAGAGTACGAGCTCCACGAGGAGCGAGTCCACCCGAG 420  
QY 159 rValValAlaHsValCyAspSerAspMetGluGluGluAspProLeuThrSerPr 179  
DB 421 TGTCTGTGACAGCTGTGTGACAGTGCATCGAGAGAGAGATCCAAAGACGTCC 480  
QY 179 oLyProLyIle 183  
DB 481 CAAGCCAAAATC 493

RESULT 8  
LOCUS BG721994 742 bp mRNA linear EST 08-May-2001  
DEFINITION 602698760F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4830759 5',  
ACCESSION BG721994  
VERSION BG721994.1 GI:14001181  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 742)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [gsabers-remail.nih.gov](mailto:gsabers-remail.nih.gov)  
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Teshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LRAM10752 row: g column: 16  
High quality sequence stop: 738.  
Location/Qualifiers

#### FEATURES

source

1..742  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4830759"  
/clone\_id="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescript (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (GcGag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size selected for average insert size 2.2 kb and  
normalized to 10<sup>7</sup> 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 180 a 202 c 193 g 167 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 1,056-77 Length: 742  
Score: 803.00 Matches: 164  
Percent Similarity: 98.26% Conservative: 5  
Best Local Similarity: 95.35% Mismatches: 3  
Query Match: 77.14% Indels: 3  
Gaps: 0

US-09-782-953-9 (1-197) x BG721994 (1-742)

QY 1 NecProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValAlaAspVal 20  
DB 228 ATGCCAGCCCTTACAGATGAGTGTGATGTTCCACTGTGCTGCTGTGTGATGATC 287  
QY 21 GluValPheThrAsnGluValLyGluLeuPheGlnGlyLeuPheArgThrTyrAsp 40  
DB 288 GAGGCTTTACCAACAGAGGTTAAGAAATTGAGGAGCTGTTCCGACTTATAT 347  
QY 41 GluCyValThrPheGlnLeuPheLeuSerPheArgArgValArgIleAsnPheSerHs 60  
DB 348 GACTGTGTGACCTTCAGCTATTTAAGATTTCAGACGTGTCCGATTAACCTTCAGCAT 407  
QY 61 ProLySerAlaAlaArgAlaArgIleGluLeuHsGluThrGlnPheArgGlyLyAslys 80  
DB 408 CCTAATCTGCAGCCCGAGCTAGATAGAGCTTCATGAAACCAATTCAAGGGGAAAAA 467  
QY 81 LeuLyLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLeuHsIleu 100  
DB 468 TTAAAGCTTACTTTCACAGGTTCAAGCTCCAGACAGATGAGAGCAAACTGCACCTT 527  
QY 101 AlaProGlnProAlaLyGlnPheLeuIleSerProPheSerProProVaGly 120  
DB 528 GGTCCAGCCCAAGCTGCAAAAGTTTCTCATCTGCCGCCCTCTCTCCACCTGTGGC 587  
QY 121 TrpLyProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaAla 140  
DB 588 TGGCAGCCCATCAACAGATGCCAGCGAGTCCCA-TATGACCTCTCTATGCTGTGGC- 645  
QY 141 LyLeuGlyProGlyGlyLeuTyrGluLeuHsAlaGlyThrGlnSerThrProSerVal 160  
DB 646 AACTAGAGCAGAGAGAGATAGAGCTTCATGAGAGGAGACTGATCAC-CCAAGTGT 704  
QY 161 ValValAlaHsValCyAspSerAspMetGluGluGln 172  
DB 705 GTCTGTGACGTTGTGACAGTGCATAGAGAGAA 740



```

RESULT 9
B0260356          546 bp      mRNA      linear      EST 06-MAY-2002
LOCUS             faal0f12.y1 zebrafish fin day3 regeneration Danio rerio cDNA clone
DEFINITION        590542 5' similar to SW:ZAK4_HUMAN Q14206 ZAK1-4 PROTEIN. , mRNA
sequence.
ACCESSION         B0260356
VERSION           B0260356.1 GI:20461126
KEYWORDS          zebrafish.
SOURCE            Danio rerio
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
REFERENCE         1 (bases 1 to 546)
AUTHORS           Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
                  ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
                  ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
                  Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
                  Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
                  and Wilson,R.
TITLE             WashU Zebrafish EST Project 1998
JOURNAL           Unpublished (1998)
COMMENT           Contact: Stephen L. Johnson
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: zbrfish@wustl.edu
                  cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
                  Matthew Clark. DNA Sequencing by: Washington University Genome
                  Sequencing Center Clone distribution: Genome Systems, St. Louis,
                  Missouri (web address: www.genomesystems.com) (email contact:
                  info@genomesystems.com) and Research Genetics, Huntsville, Alabama
                  (web address: www.resgen.com) (email contact: info@resgen.com) and
                  RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
                  www.rzpd.de)
FEATURES          Seq primer: T3 ET from Amersham
                  High quality sequence stop: 395.
source            Location/Qualifiers
                  1..546
                   /organism="Danio rerio"
                   /db_xref="taxon:7955"
                   /clone="5903542"
                   /clone_lib="zebrafish fin day3 regeneration"
                   /sex="mixed male and female"
                   /tissue_type="3 day fin regenerates"
                   /lab_host="E. coli XL0LR"
                   /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; 1st
strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed
by second strand synthesis, and ligated to 5' adapter (5'
)-aattcgacagag-3', 3'-gccgtgtctc-5'. cDNA was cloned
directionally (EcoRI/XhoI) into Stratagene Zap express
lambda phage arms. Mass invivo excision done to obtain
inserts in pBK-CMV phagemid."
BASE COUNT        144 a 151 c 139 g 112 t
ORIGIN
Alignment Scores:
Pred. No.:        3,1e-76          Length:      546
Score:            788.00          Matches:     148
Percent Similarity: 89.56%        Conservative: 15
Best Local Similarity: 81.32%      Mismatches:  19
Query Match:      75.70%          Indels:      0
Deletions:        14              Gaps:        0
US-09-782-953-9 (1-197) x BQ260356 (1-546)
QY 13 LeuValAlaCysValValAlaPheThrAsnGlnGluValLysGluTyrPhe 32
Db 1 CTCTGCCTGTCGGTGGACGTGGAGGTGTTCCACGATCGAGAGGTCAAGGAAGTTT 60
QY 33 GluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSerPheArg 52

```

```

Db 61 GAAGCGCTCTTCTCGCTACGATGAGGTGTGACGTTCCAGCAGTTTAAGAGCTTCAGA 120
QY 53 ArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGluLeuHis 72
Db 121 CGTGTACGAATCAATTTAGTACCTTAAGCAGCCAGCGCAAGATAAGAGCTGCAC 180
QY 73 GluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThrProGlu 92
Db 181 GAGACGCTTTTCAGAGGGAAGAACTCAAACTTTTACTTTGCCAGGTGCAGAACCCAGTA 240
QY 93 ThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuLeuSer 112
Db 241 TCTGAAGGAGACAATTTGCACCTGGCTCTCCACAGCCTTCCAAACAGTTCCTCATCTCT 300
QY 113 ProProSerSerProProValGlyTyrLysProLysSerAspAlaThrProValLeuAsn 132
Db 301 CCCCTGGCTCACCCTCCAGTCGGTTGGCAACAGATTGATGAGCCACACCGGTTATTAAAC 360
QY 133 TyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeuHisAla 152
Db 361 TACGACCTGCTGTATGTCAGTTGCCAAGCTTGGCCAGGTGAGAAGTTCGAGTCCATGCG 420
QY 153 GlyThrGluSerThrProSerValValHisValCysAspSerAspMetGluGluGlu 172
Db 421 GGCACAGAGTCTACGCCCCAGTCGTGGTGTCCACGCTCTGTGACAGCGACACGATGAGGAG 480
QY 173 GluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192
Db 481 GAGGACCAACCAAGACACACCTTAACCAAAATAATCCAGACCCGACGCTCCAGACCTGCCA 540
QY 193 ProSer 194
Db 541 GTCTCA 546
RESULT 10
B1476888
LOCUS             B1476888.1 GI:15310227
DEFINITION        ; mRNA sequence.
ACCESSION         B1476888
VERSION           B1476888.1
KEYWORDS          EST.
SOURCE            zebrafish.
ORGANISM          Danio rerio
                  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
REFERENCE         1 (bases 1 to 617)
AUTHORS           Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
                  ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
                  ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
                  Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
                  Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
                  and Wilson,R.
TITLE             WashU Zebrafish EST Project 1998
JOURNAL           Unpublished (1998)
COMMENT           Contact: Stephen L. Johnson
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: zbrfish@wustl.edu
                  cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
                  Matthew Clark. DNA Sequencing by: Washington University Genome
                  Sequencing Center Clone distribution: Genome Systems, St. Louis,
                  Missouri (web address: www.genomesystems.com) (email contact:
                  info@genomesystems.com) and Research Genetics, Huntsville, Alabama
                  (web address: www.resgen.com) (email contact: info@resgen.com) and
                  RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
                  www.rzpd.de)
                  High quality sequence stop: 467.

```



to Rot = 20.0 and subtraction to Rot = 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCTGAGTTAATAATTAATCCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I."

BASE COUNT 149 a 183 c 158 g 150 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,3e-74 Length: 640  
Score: 774.00 Matches: 147  
Percent Similarity: 99.32% Conservatives: 0  
Best Local Similarity: 99.33% Mismatches: 1  
Query Match: 74.35% Indels: 0  
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x BF534085 (1-640)

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
DB 197 ATGCCAGCCCTAGCATGAGCTGTGATGTTTCCACTCTGTCGCTGTGTGGTGTG 256  
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluLysPheArgThrTyRAsp 40  
DB 257 GAGGTCTTTACCAATCAGAGGTTAAGGAAATTCGAGGACTTTCGGACCTATGAT 316  
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60  
DB 317 GAATGTGTGAGCTTCCAGCTGTTTAAGAGTTTCCGAGGGTTTCGATAAATTCAGCCAT 376  
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80  
DB 377 CCCAAATCTGCAGCCCGTCCCGATAGAGCTTCATGAGACTCAGTTCCAGAGGGAAG 436  
QY 81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspClyAspLysLeuHisLeu 100  
DB 437 CTAAACTCTACTTCCGCCAGGTCAGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAG 496  
QY 101 AlaProGlnProAlaLysGlnPheLeuIleSerProProSerSerProProValGly 120  
DB 497 GCACCTCCAGCCCTGCCAAGAGTTCCTCACTCACTCCCTCTCTCTCTCTCTCTCT 556  
QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyRAspLeuLeuTyRAlaValAla 140  
DB 557 TGGAGCCCTATCAGCGATGCCACACAGTCTCACTCACTATGACCTCTTTATGCTGTGCC 616  
QY 141 LysLeuGlyProGlyGlyLysTyR 148  
DB 617 AAATAGGACCGAGGAGAAATAT 640

BF534085 1007 bp mRNA linear EST 11-DEC-2000  
LOCUS 602075073F1 NCI\_CGAP\_Li9 Mus musculus cDNA clone IMAGE:4211994 5',  
DEFINITION mRNA sequence.

ACCESSION BF534085  
VERSION BF534085.1 GI:11621448

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 1007)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM9782 row: a column: 19

High quality sequence stop: 742.

## FEATURES

Location/Qualifiers

1..1007  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4211994"  
/clone\_lib="NCI CGAP\_Li9"  
/lab\_host="DH10B (TI phase-resistant)"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 249 a 283 c 255 g 220 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.16e-72 Length: 1007  
Score: 758.50 Matches: 159  
Percent Similarity: 86.46% Conservatives: 7  
Best Local Similarity: 82.81% Mismatches: 20  
Query Match: 72.86% Indels: 6  
DB: 12 Gaps: 1

US-09-782-953-9 (1-197) x BF534085 (1-1007)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
DB 264 AACTCATCTGTTGCTGCAATGTTCCACAGTCTGTGTTTGAAGAGAGAGAGAGAGAA 323  
QY 31 LysPheGluGlyLeuPheArgThrTyRAspGluCysValThrPheGlnLeuPheLysSer 50  
DB 324 AAATTCGAGGGGACTTTCGGACCTATGATGAATGTGTGACGTTTCAAGCTT 383  
QY 51 PheArgAlaGlyAlaArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
DB 384 TTCCAGCGGTTTCGATAAATTTTCAGCCATCCCAATCTGCAGCCCGTCCCGGATAGAG 443  
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyRLeuAlaGlnValGlnThr 90  
DB 444 CTTCATGAGACTCAGTTCCAGAGGGAAGAGCTAAACTCTACTTCCGCCAGGTCCAGACC 503  
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110  
DB 504 CCAGAGACAGATGGAGACAACTGCATTTGGCACCTCCACAGCCTGCCAAACAGTTCTCT 563  
QY 111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130  
DB 564 ATCTCACCCCTTCATCT 623  
QY 131 LeuAsnTyRAspLeuLeuTyRAlaValAlaLysLeuGlyProGlyGlyLysTyRGlul 150  
DB 624 CTCAACTATGACCTCTTTATGCTGTGGCCAACTAGACACAGGAGCAGAGTAGTATGAC 683  
QY 150 euHisAlaGlyThrGlnSerThrProSerValValHisValCysAspSerAspMe 169  
DB 684 TGCAGCTGGAACTAGTCTACACCGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743  
QY 169 tGluGluGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgPr 189  
DB 744 GCGCGCGGAAGAACCCCAAGAAATT---CCCAAGCCAAAATCATTTCCAGACCGGAGTCC 800  
QY 189 OGlyLeu-ProProSerValSerAsn 197  
DB 801 TGGCTGGTCAACCATCGGGTCCAAT 826  
RESULT 13  
LOCUS AI796134/c 557 bp mRNA linear EST 06-JUL-1999

DEFINITION	wh42909.x1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:2383456 3'					
ACCESSION	A1796134					
VERSION	A1796134.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Euryarchaea; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 557) NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)					
JOURNAL COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbb-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://www.bio.lnl.gov/bbrp/image/image.html">www.bio.lnl.gov/bbrp/image/image.html</a> Seq primer: -40UP from Gibco High quality sequence group: 455. Location/Qualifiers 1..557 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2383456" /clone_lib="NCI_CGAP_Kid11" /lab_host="DH10B" /note="Organ: kidney; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and as circles were used in vitro. Following HATP hybridization reaction, this DNA was used as tracer in a subtractive library construction. The driver was PCR-amplified cDNA from a pool of 5,000 clones made from the same library (clonoids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."					
BASE COUNT	125 a	107 c	162 g	162 t	1 others	
ORIGIN						
Alignment Scores:						
Pred. No.:	7,61e-72	Length:	557			
Score:	748.00	Matches:	141			
Percent Similarity:	97.3%	Conservative:	6			
Best Local Similarity:	93.3%	Mismatch:	4			
Query Match:	71.85%	Indels:	0			
DB:	9	Gaps:	0			
DS-09-782-953-9 (1-197) x A1796134 (1-557)						
Oy	46	GInLeuPhelYSeSPheARgVAlArGiLeAsnPhSeRhiABroLYSSeRaLaAa	65			
Dd	557	CAGCATTAAAGATTCAACAGTCGTCCGTATAACTTCAGCATCTTAATCTGCAGCC	498			
Oy	66	ArgAlaargIGlGLeUuhiegluthrGlnPhleatrgILylyLeuleulyLeuTYrPhe	85			
Dd	497	CGAGCTAGAGATGAGCTTCATCATGAACCACCAATTCAGGGGAAAAATTTAAAGCTCTACTT	438			
Oy	86	AlAGInValGlThPrOGluThrPaRGiYaERLyLeuHiSlenuLabProPGInPRo	105			
Dd	437	GCACAGGTCACACTCCACAAGACAGATGAGAACAATCACAATCTGCCTCACACCCAGCCT	378			
Oy	106	AlAtYgInPhLeuileiseerPropofoseSerPropRoVaIgLYrPyPsProlieser	125			
Dd	377	GCCAAACAGTTTTCATCTCGCCCCCTTCTCCGCCACCTGTTGGCTGGCAGACCCATCAAC	318			
Oy	126	ApaLatMrProValleuanTyArEpLeuLeuTYraLeValAlalyLeuGlYProGIY	145			

Db	317	GATGCCAGCCAGTCTCTCACTATGACCTCTCTATATGCTGTGGCCAAACTAGAGCAGAGA	258
Qy	146	GlutylTyrGluLeuHisAlaGlyThrIleSerThrProSerValValHisValCys	165
Db	257	AAAAAGTATGAGCTCTCAGCAGGAGGACTGATGCCACCCCAAGATGTCGTGCACCTGTCC	198
Qy	166	AppSerAspMetGluGluGluAspProIleValThrSerProIleValIleLeuGln	185
Db	197	GACGCTGCAATAGAGAAAGAGAGAGCCCAAGACTTCCCAAGGCCAAATATCATCCAA	138
Qy	186	ThrArgArgProGlyLeuProProSerValSer	196
Db	137	ACTCGGGTCTCTGCGCTCCACCTCCCTGTCTC	105
RESULT 14			
LOCUS	BG701025	626 bp	mRNA linear EST 07-MAY-2001
DEFINITION	602682030P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814743 5',		
ACCESSION	BG701025		
VERSION	BG701025.1	GI:13970954	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 626)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>		
	Tissue Procurement: Mikhail Palkovite, M.D., Ph.D.		
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	<a href="http://image.lnl.gov">http://image.lnl.gov</a>		
	Plate: LRAM10710 row: 1 column: 08		
	High quality sequence stop: 626.		
FEATURES			
source	Location/Qualifiers		
	1..626		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4814743"		
	/clone_lib="NIH_MGC_95"		
	/class_type="hippocampus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptpr (modified		
	pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Grogas		
	) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',		
	size-selected for average insert size 2.5 kb and		
	normalized to R0F 5. This is a primary library enriched		
	for full-length clones and constructed using the		
	Cap-trapper method (Carninci, in preparation). Library		
	constructed by M. Brownstein (NIH/NHGRI, National		
	Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	150 a 177 c 154 g 145 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..37e-69	Length:	626
Score:	728.00	Matches:	138
Percent Similarity:	99.50%	Conservative:	4
Best Local Similarity:	96.50%	Mismatches:	1
Query Match:	69.93%	Indels:	0
DB:	12	Gaps:	0

```

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 196 ATGCCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGGTTGGCTGTGTGGATGTC 255
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40
Db 256 GAGGTCTTTTACCATCAGAGAGTTAAGGAAAATTTGAGGAGCTGTTCCGGACTTATGAT 315
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgIleAsnPheSerHis 60
Db 316 GACTGTGTGAGCTTCCAGCTATTTAAAGAGTTTCAGACGTGTGCCGTATATAAATTCAGCAAT 375
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 376 CCTAATCTGAGCCCGAGCTAGGATAGAGCTTCATGGAACCCCAATTCAGAGGGGAAAAA 435
QY 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 436 TTAAGGCTCTACTTTGCACAGTTTCAGACTCCAGACAGATGGAGACAACTGCATTTG 495
QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 496 GCTCCACCCAGCCTGCAACAGTTTCTCATCTCGCCCTCTCTCCACCTGTTGGC 555
QY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaValAla 140
Db 556 TGGCAGCCCAACCAAGTGCACGCCAGTCTCACTTACCTATGACCTCTATGCTGTGGCC 615
QY 141 LysLeuGly 143
Db 616 AAAC TAGGA 624

```

```

RESULT 15
AU169846 747 bp mRNA linear EST 29-JAN-2001
LOCUS AU169846 Ol-br-ad cDNA Oryzias latipes cDNA clone br5301, mRNA
DEFINITION

```

```

ACCESSION AU169846
VERSION AU169846
KEYWORDS AU169846.1 GI:12591915
SOURCE EST.
ORGANISM Japanese medaka.
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 747)
Mita, K., Iehikawa, Y. and Yamauchi, M.
Establishment of cDNA database of medaka, Oryzias latipes
Unpublished (2001)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
method: uni-directional sequence direction: sequenced from T3 primer
(5' -> 3').

```

```

FEATURES
source Location/Qualifiers
1..747
/organism="Oryzias latipes"
/strain="HNI"
/db_xref="taxon:8090"
/clone="br5301"
/clone_lib="Ol-br-ad cDNA"
/sex="female/male mixed"
/tissue type="brain"
/dev_stage="adult"
BASE COUNT 193 a 210 c 191 g 153 t
ORIGIN

```

```

Alignment Scores: 2.85e-69 Length: 747
Pred. No.: 726.00 Matches: 138
Score:

```

```

Percent Similarity: 82.74% Conservative: 25
Best Local Similarity: 70.05% Mismatches: 32
Query Match: 69.74% Indels: 2
DB: 9 Gaps: 1

```

US-09-782-953-9 (1-197) x AU169846 (1-747)

```

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 155 ATGCCATCGATAAATATGAACCTGGAGATGGTGCACCATCATCGCCGCGAAGTAGACACC 214
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40
Db 215 GAAGTGTCTTACGTTCAGACAAATATGCAGGAGAGGTTTGAAGCCGTGTTGAGGATTTATGAT 274
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgIleAsnPheSerHis 60
Db 275 GAACACAAACCTTCCAGATGTTCAAAAGCTTCAGAAGAGTGGGATCAACTTCAGCACC 334
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 335 CCAGAGGCGAGCCCGCGCTGCATCGAGTTGCACGAATCGAGTTCACGCGCAAGAG 394
QY 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 395 CTCAAACTCTACTTTGCTCAAATCCAGAAATGGTGACGAGACATAGACAAAGTCTTACCTG 454
QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 455 GCGCCCTCCCAACCCCGTGAACACAGTTTCTGATCTCTCTCTGCTCGCGCGCTGTGGGA 514
QY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaValAla 140
Db 515 TGGAGCGCAGATGAAGACGCCACCCGTCATCACTACGACCTGCTATGCTGTTGCC 574
QY 141 LysLeuGlyProGlyGlyLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 575 AAGCTCGGACCTGGTGAAAGATATGAACCTCCACGCGGAAACAGAGTCCCACTCCAAGTGT 634
QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
Db 635 GTCCGTCACGTTGGATAGC-----GAGGCGAGGAGGACGAGCGGTGCGGCCAAAG 688
QY 181 ProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 689 CAGCAGATGCTCCAGACCAAGCGGTCCCGATGTTCCCGAACCGGTCTCCAAC 739

```

Search completed: December 14, 2002, 22:28:15  
Job time : 1116.9 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.5632 seconds  
(without alignments)  
858.887 Million cell updates/sec

Title: US-09-782-953-9  
Perfect score: 1041  
Sequence: 1 MPAPSMDCVSTLVACVDV.....SPKPKIOTRRPGLPFSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	23	AAE18913
2	980	94.1	192	23	AAE18916
3	905	86.9	234	22	AAE18917
4	905	86.9	255	23	AAE18918
5	658.5	63.3	241	23	AAE18917
6	641	61.6	197	23	AAE18914
7	633.5	60.9	212	23	AAE18918
8	632.5	60.8	198	23	AAE18912
9	622	59.8	197	23	AAE18915
10	613.5	58.9	198	23	AAE18911
					Mouse MCIP associa
					Human MCIP associa
					Amino acid sequenc
					Human Down syndrom
					Human MCIP associa
					Human MCIP associa
					Human MCIP associa
					Mouse MCIP associa
					Human MCIP associa

11	586	56.3	171	20	AAW73898
12	404	38.8	292	22	ABB71467
13	323	31.0	111	21	AAE18913
14	254.5	24.4	142	21	AAE18913
15	219	21.0	58	22	ABB29495
16	219	21.0	58	22	ABB34672
17	219	21.0	58	22	ABB36877
18	219	21.0	58	22	ABB20081
19	219	21.0	58	22	AAE18913
20	219	21.0	58	22	AAE18913
21	219	21.0	58	22	AAE18913
22	219	21.0	58	22	AAE18913
23	219	21.0	58	22	AAE18913
24	219	21.0	58	22	AAE18913
25	219	21.0	58	22	AAE18913
26	219	21.0	58	22	AAE18913
27	219	21.0	58	22	AAE18913
28	219	21.0	58	22	AAE18913
29	219	21.0	58	22	AAE18913
30	174	16.7	56	22	ABB28178
31	174	16.7	56	22	ABB33353
32	174	16.7	56	22	ABB18813
33	174	16.7	56	22	AAE18913
34	174	16.7	56	22	AAE18913
35	174	16.7	56	22	AAE18913
36	174	16.7	56	22	AAE18913
37	174	16.7	56	22	AAE18913
38	174	16.7	56	22	AAE18913
39	133	12.8	25	22	AAE18913
40	101	9.7	454	17	AAE18913
41	101	9.7	454	17	AAE18913
42	90.5	8.7	434	21	AAE18913
43	89	8.5	777	23	ABP41839
44	89	8.5	874	22	ABP70570
45	87	8.4	732	23	ABP93616

## ALIGNMENTS

RESULT 1  
AAE18913  
ID AAE18913 standard; Protein; 197 AA.

AC AAE18913;  
DT 17-MAY-2002. (first entry)  
XX Mouse MCIP associated proein #4.  
DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; mouse.  
XX Mus musculus.  
XX WO200204491-A2.  
XX 17-JAN-2002.  
XX 06-JUL-2001; 2001WO-US21662.  
XX 07-JUL-2000; 2000US-216601P.  
XX 13-FEB-2001; 2001US-0782953.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX (WILL/) WILLIAMS S R.  
XX (ROTH/) ROTHERMEL B.  
XX Williams SR, Rothermel B;  
XX WPI; 2002-179698/23.  
XX N-P5DB; AAD30154.

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

XX Disclosure: Page 154, 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated protein.  
 CC Note: This sequence has been described as mouse MCIP2 encoding DNA in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.

XX Sequence 197 AA;

Query Match 100.0%; Score 1041; DB 23; Length 197;

Best Local Similarity 100.0%; Pred. No. 7e-104;  
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAPSDCVSTLVACVNDVEVTNOEVEKEFGLFRITDECTFOLFSPFRVRINFESH 60  
 DB 1 MPAPSDCVSTLVACVNDVEVTNOEVEKEFGLFRITDECTFOLFSPFRVRINFESH 60  
 QY 61 PXSAAARIEIHEHTOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVG 120  
 DB 61 PXSAAARIEIHEHTOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVG 120  
 QY 121 WKRISSDATPVLYNDLLVYAVAKLGPGEKYEIHLACTESTPSVVVAVCDSDMEEDDPTSPK 180  
 DB 121 WKRISSDATPVLYNDLLVYAVAKLGPGEKYEIHLACTESTPSVVVAVCDSDMEEDDPTSPK 180  
 QY 181 PKIITRRRRLPPSVSN 197  
 DB 181 PKIITRRRRLPPSVSN 197

RESULT 2  
 AAE18916  
 ID AAE18916 standard; Protein, 192 AA.  
 AC AAE18916;  
 XX  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #3.  
 XX  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;  
 XX WPI: 2002-179698/23.  
 DR N-PSDB; AAD30157.  
 XX

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

XX Example 1: Page 165-166; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as human MCIP splice variant in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.

XX Sequence 192 AA;

Query Match 94.1%; Score 980; DB 23; Length 192;

Best Local Similarity 96.4%; Pred. No. 2.5e-97;  
 Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDSTLVACVNDVEVTNOEVEKEFGLFRITDECTFOLFSPFRVRINFESH 65  
 DB 1 MDCDSTLVACVNDVEVTNOEVEKEFGLFRITDECTFOLFSPFRVRINFESH 65  
 QY 66 RARIELHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVWQPIN 125  
 DB 61 RARIELHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVWQPIN 120  
 QY 126 DATPVLYNDLLVYAVAKLGPGEKYEIHLACTESTPSVVVAVCDSDMEEDDPTSPK 185  
 DB 121 DATPVLYNDLLVYAVAKLGPGEKYEIHLACTESTPSVVVAVCDSDMEEDDPTSPK 180  
 QY 186 TRRPGRLPPSVSN 197  
 DB 181 TRRPGRLPPSVSN 192

RESULT 3  
 AAB31788  
 ID AAB31788 standard; Protein, 234 AA.

AC AAB31788;  
 XX

DT 30-APR-2001 (first entry)  
 DE

XX Amino acid sequence of a human detoxification protein.

XX Human; detoxification protein; DETX; cancer; leukemia; melanoma;  
 KW adenocarcinoma; autoimmune disorder; inflammatory disorder;  
 KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
 KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
 KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis.

XX Homo sapiens.  
 XX

OS Key Location/Qualifiers  
 FH Peptide 1-27  
 FT /note= "signal peptide"

PT





```

QY 11 STLVACVVDVBEFTNQEVKEKEFEGLFRTYDECVTQFLFKSPFRVRINFSHPKSAARARIE 70
DB 69 NSLFACVNHQSVFEGESENKFEGLFRITDDCVTPQLKSPRRVRINFSHPKSAARARIE 128
OY 71 LHETQFRGKKLKYFAOVOTPETDDKHLAPPOKQFLISPPSSPVGMKPSIDATPV 130
DB 129 LHETQFRGKKLKYFAOVOTPETDDKHLAPPOKQFLISPPSSPVGMKPSIDATPV 188
OY 131 LNYDLLYAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDPKTSPPKIIQTRRPG 190
DB 189 LNYDLLYAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDPKTSPPKIIQTRRPG 248
OY 191 LPPSVSN 197
DB 249 LPPSVSN 255

RESULT 5
AAE18917
ID AAE18917 standard; Protein; 241 AA.
XX
AC AAE18917;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated protein #4.
XX
KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; human.
XX
OS Homo sapiens.
XX
PN MO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001MO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
XX PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (MILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.
XX
PI Williams SR, Rothermel B;
XX
DR WPI; 2002-179698/23.
XX DR N-PSDB; AAD30158.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Example 1; Page 168-169; 174pp; English.
XX
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX complex with the catalytic subunit of calcineurin and increased levels
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX transcription of certain target genes. The invention also relates to
XX methods for identifying modulators of MCIP binding, expression or
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX may be used for treating cardiac hypertrophy and heart failure.
XX Antisense to MCIP can be used in characterizing for determining the
XX healthy and diseased tissues and subsequently for determining the
XX presence or absence of cardiomyopathy or as predictor of heart disease.
XX The present sequence is human MCIP associated protein.
XX Note: This sequence has been described as splice variant of MCIP1
XX initiated by exon 4 in the specification, however the sequence seems
XX to be a MCIP associated protein.

```

```

SQ Sequence 241 AA;
Query Match 63.3%; Score 658.5; DB 23; Length 241;
Best Local Similarity 70.1%; Pred. No. 1,6e-62;
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

QY 11 STLVACVVDVBEFTNQEVKEKEFEGLFRTYDECVTQFLFKSPFRVRINFSHPKSAARARIE 70
DB 47 TSLPACSVHEAVFEARQKERFEALFTYDDQVTFQFLKSPFRVRINFSHPKSAARARIE 106
OY 71 LHETQFRGKKLKYFAOVOTPETDDKHLAPPOKQFLISPPSSPVGMKPSIDATPV 130
DB 107 LHETDFNGQKKLKYFAOVQSGEVKDSYLLPPOKQFLISPPSPVGMKPSIDATPV 166
OY 131 LNYDLLYAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDPKTSPPKIIQTRRPG 190
DB 167 LNYDLLCAVSKLGPGEKYEELHAGTSTPSVVVHVCSESETEEBETK-NPQKIAQTRPD 225
OY 191 LPPSVSN 197
DB 226 PPTAALN 232

RESULT 6
AAE18914
ID AAE18914 standard; Protein; 197 AA.
XX
AC AAE18914;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated protein #1.
XX
KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; human.
XX
OS Homo sapiens.
XX
PN MO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001MO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
XX PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (MILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.
XX
PI Williams SR, Rothermel B;
XX
DR WPI; 2002-179698/23.
XX DR N-PSDB; AAD30155.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Example 1; Page 157-158; 174pp; English.
XX
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX complex with the catalytic subunit of calcineurin and increased levels
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX transcription of certain target genes. The invention also relates to
XX methods for identifying modulators of MCIP binding, expression or
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX may be used for treating cardiac hypertrophy and heart failure.
XX Antisense to MCIP can be used in characterizing for determining the
XX healthy and diseased tissues and subsequently for determining the

```

CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as murine splice variant in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 XX  
 SQ Sequence 197 AA;

Query Match 61.6%; Score 641; DB 23; Length 197;  
 Best Local Similarity 64.2%; Pred. No. 9.2e-61;  
 Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

QY 11 STLAVCVVVEFTNOEVEKEFLPTDYECVTFQFKFRRVRINFSHPKSAARARIE 70  
 DB 11 SLLIACVANSDFSESTRAKFESLFTYDKDITFOYFKSPKRVIRINFSNPFSAADARLQ 70  
 QY 71 LHETQFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKPISDATPV 130  
 DB 71 LHKTEFLGKMKLYFAOTLHIGSS----HLAPNPDKOFLISPPSPVGVKQVEDATPV 126  
 QY 131 LNYDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEEDPKTS----PKPKIIQT 186  
 DB 127 INYDLLYAIKSLGPGKYLHAGTSTPSVVVHVCDSDMEEDPKTS----PKPKIIQT 186  
 QY 187 RRCLPP 193  
 DB 187 RRPETP 193

RESULT 7  
 AAE18918  
 ID AAE18918 standard; Protein; 212 AA.  
 AC AAE18918;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated protein #5.  
 XX  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 DR WPI; 2002-179698/23.  
 DR N-PSDB; AAD30159.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 171-172; 174pp; English.  
 XX

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate

CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterising the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein described in the  
 CC invention.  
 XX  
 SQ Sequence 212 AA;

Query Match 60.9%; Score 633.5; DB 23; Length 212;  
 Best Local Similarity 67.9%; Pred. No. 6.5e-60;  
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

QY 11 STLAVCVVVEFTNOEVEKEFLPTDYECVTFQFKFRRVRINFSHPKSAARARIE 70  
 DB 28 TSLFACSVHEAVFEAREQERFEALFTYDDQVTFQFKSPRRVRINFSKPEAARARIE 87  
 QY 71 LHETQFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKPISDATPV 130  
 DB 88 LHETDNGKLLKLYFAQ-----SYLLPPQPVKQFLISPPSPVGVKQSEDAMPV 137  
 QY 131 LNYDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEEDPKTSKPCKIIQTRRPG 190  
 DB 138 INYDLLCAVSKLGPGEKYLHAGTSTPSVVVHVCDSETEEEETK-NPKQKIAQTRPD 196  
 QY 191 LPPSVSN 197  
 DB 197 PPTAALN 203

RESULT 8  
 AAE18912  
 ID AAE18912 standard; Protein; 198 AA.  
 AC AAE18912;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Mouse MCIP associated protein #3.  
 XX  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; mouse.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 DR WPI; 2002-179698/23.  
 DR N-PSDB; AAD30153.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 151-152; 174pp; English.  
 XX

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC Note: This sequence has been described as human MCIP3 in the  
 CC specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 CC Sequence 198 AA;  
 SQ  
 Query Match 60.8%; Score 632.5; DB 23; Length 198;  
 Best Local Similarity 64.4%; Pred. No. 7,66-60;  
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;  
 QY 11 STLVACVADVVEPTNOEVEKEFEGLFRTYDECVTPLKSPFRVRINFSHPSAARIE 70  
 DB 11 SSLICVANDVFSSESTRAKESLFRYDKDTTFQYFSPFRVRINFSNPLSADARLR 70  
 QY 71 LHETQFRGKKLKYFAOVQTPETDGDKHLAPPOPAKQFLISPPSPVGVKPISDATPV 130  
 DB 71 LKTEFLGKEMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVKQVEDATPV 126  
 QY 131 LNYDLLVAVALKPGCEKELYLAAGTESTPSVVHVCDSDMEEDPK-----TSPKRIIQ 185  
 DB 127 INYDLLVAISKPGCEKELYLAATDTPSVVVHVCDSDENEBEEMERMRPKRIIQ 186  
 QY 186 TRRPGLR 193  
 DB 187 TRPEYTP 194  
 RESULT 9  
 AAE18915  
 ID AAE18915 standard; Protein; 197 AA.  
 AC AAE18915;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #2.  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human.  
 OS Homo sapiens.  
 PN WO200204491-A2.  
 PD 17-JAN-2002.  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 PI Williams SR, Rothermel B;  
 DR MPI: 2002-179698/23.  
 DR N-PSDB; AAD30156.  
 XX Screening for modulators of muscle calcineurin interacting protein

PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 XX Example 1; Page 161-162; 174pp; English.  
 XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC Note: This sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 CC Sequence 197 AA;  
 SQ  
 Query Match 59.8%; Score 622; DB 23; Length 197;  
 Best Local Similarity 63.6%; Pred. No. 1e-58;  
 Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;  
 QY 11 STLVACVADVVEPTNOEVEKEFEGLFRTYDECVTPLKSPFRVRINFSHPSAARIE 70  
 DB 11 SATIACHDPRVFDGLORAFESLFRYDKDTTFQYFSPFRVRINFSNPLSADARLRQ 70  
 QY 71 LHETQFRGKKLKYFAOVQTPETDGDKHLAPPOPAKQFLISPPSPVGVKPISDATPV 130  
 DB 71 LKTEFLGKEMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVKQVEDATPV 126  
 QY 131 LNYDLLVAVALKPGCEKELYLAAGTESTPSVVHVCDSDMEEDPKS---PKRIIQ 186  
 DB 127 INYDLLVAISKPGCEKELYLAATDTPSVVVHVCDSDENEBEEMERMRPKRIIQ 186  
 QY 187 RRPGLRP 193  
 DB 187 RRPETYP 193  
 RESULT 10  
 AAE18911  
 ID AAE18911 standard; Protein; 198 AA.  
 AC AAE18911;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #2.  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human.  
 OS Homo sapiens.  
 PN WO200204491-A2.  
 PD 17-JAN-2002.  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX

```

PI Williams SR, Rothermel B;
XX WPI; 2002-179698/23.
DR N-PSDB; AAD30152.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
XX Disclosure; Page 148-149; 174pp; English.
PS
XX
CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC transcription of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterising the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is human MCIP associated protein.
XX
SQ Sequence 198 AA;
Query Match 58.9%; Score 613.5; DB 23; Length 198;
Best Local Similarity 63.3%; Pred. No. 8.5e-58;
Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;
QY 11 STLAVCVVVEVTNQEVKFKFGLFRTYDECVTFOLFKSFRVRINFSPKSAARIE 70
DB 11 SATIACHLDRVFDGLCRAKFESLFRTYDKDTTFQYFKSFRVRINFSPKSAARLR 70
QY 71 LHETQFGKGLKLYFAQVQTPETDGDKLHAPPQAPAKQFLISPPSPVGVKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVKQVEDATPV 126
QY 131 LNYDLLYAVAKLGFGKGYELHAGTSTPSVVHVCDSDMEEDDPK----TSPKPKIIQ 185
DB 127 INYDLLYAIKLGFGKGYELHAAATDTPPSVVHVCDSDQNEEEEMERMRKPKKIIQ 186
QY 186 TRRPGLPP 193
DB 187 TRRPEYTP 194
RESULT 11
ID AAW73898 standard; Protein; 171 AA.
XX
AC AAW73898;
XX
DT 09-APR-1999 (first entry)
XX
DE Human Down's Syndrome critical region 1 protein.
XX
KW DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
KW Central Nervous System development; mental retardation; heart defect.
XX
OS Homo sapiens.
XX
PN US5869318-A.
XX
PD 09-FEB-1999.
XX
PF 07-JUN-1996; 96US-0665040.
XX
PR 07-JUN-1995; 95ES-0001140.
XX
PA (PALL/) PALLEJA X E.
XX
PI Fuentes JJ, Palleja XE, Pritchard M;
XX WPI; 1999-152781/13.
DR N-PSDB; AAX01282.
XX
PT DNA encoding foetal brain proteins - believed to be associated with
PT Down's syndrome
XX
XX Claim 4; Column 15-18; 19pp; English.
PS
XX
CC This sequence is encoded by the Down's Syndrome critical region 1 (DSCR1)
CC gene of the invention. The DSCR1 gene was found to be located in the
CC q22.1-22.2 region of human chromosome 21. An increase in the transient
CC expression of DSCR1 mRNA in the brains of young rats, compared to
CC expression levels in the brains of adult rats, suggests an important role
CC for DSCR1 during the development of the Central Nervous System (CNS), and
CC that over expression of DSCR1 may be involved in pathogenic abnormalities
CC of mental retardation and/or heart defects as found in Down's syndrome
CC patients.
XX
SQ Sequence 171 AA;
Query Match 56.3%; Score 586; DB 20; Length 171;
Best Local Similarity 66.5%; Pred. No. 6.4e-55;
Matches 113; Conservative 22; Mismatches 27; Indels 8; Gaps 2;
QY 28 VKEPEGLFRTYDECVTFOLFKSFRVRINFSPKSAARIEHETQFGKGLKLYFAQ 87
DB 2 VYAKFESLFRTYDRDITTFQYFKSFRVRINFSPKSAARIEHETQFGKGLKLYFAQ 61
QY 88 VQTPETDGDKLHAPPQAPAKQFLISPPSPVGVKQVEDATPVNLYDLLYAVAKLGPGEK 147
DB 62 TLHIGSS----HLAPPNDKQFLISPPSPVGVKQVEDATPVNLYDLLYAVAKLGPGEK 117
QY 148 YELHAGTSTPSVVHVCDSDMEEDDPKTS----PKKIIQTRRPGLPP 193
DB 118 YELHAATDTPSVVVHVCDSDQNEEEEMERMRMRPKKIIQTRRPEYTP 167
RESULT 12
ID ABB71467 standard; Protein; 292 AA.
XX
AC ABB71467;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41193.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL15570.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

```

XX Disclousure; SEQ ID NO 41193; 21pp + Sequence Listing; English.  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA  
 CC sequences (AB161840-AB16175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WFO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 292 AA;  
 SO  
 Query Match 38.8%; Score 404; DB 22; Length 292;  
 Best Local Similarity 43.2%; Pred. No. 5.6e-35;  
 Matches 83; Conservative 35; Mismatches 62; Indels 12; Gaps 4;  
 Oy 4 PSMDD-----STLVACVDEVEFTNOEVEKEFEGLFRTYDECVTFOLFKSFRVRINF 58  
 Db 104 PEVDADSFDDLTPTSLITVINHSEVFANPELKIAMEELFTTSBSATFQWLRSFRRLRVY 163  
 Oy 59 SHPKSARARIELHETQFRGKK-LKLYFAOVQTPETDGDKLHLAPQAPQFLISPPSP 117  
 Db 164 DNAIAAANRKLHGYEFKKTIVITCFAPQPTPVSN---XNLQPPAPVKQFLISPPAP 220  
 Oy 118 PVGWRKISDAPLVNLDLYAVAKLGPGEKYLHAGTSTPSVVVHVCSDMBEEDPPT 177  
 Db 221 PAGMEPREGEPLVNHDLALALSLTGESHEILHPSQSDQPAIIVHTA---MLAETGPG 277  
 Oy 178 SPKPKIOTRRP 189  
 Oy 278 QVKAPIVOTKCP 289  
 Db  
 RESULT 13  
 AAG01768  
 ID AAG01768 standard; Protein; 111 AA.  
 XX  
 XX AAG01768;  
 AC  
 XX  
 XX 06-OCT-2000 (first entry)  
 DT  
 XX  
 XX Human secreted protein, SEQ ID NO: 5849.  
 DE  
 XX  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP1033401-A2.  
 PN  
 XX  
 XX 06-SEP-2000.  
 PD  
 XX  
 XX 21-FEB-2000; 2000EP-0200610.  
 PP  
 XX  
 XX 26-FEB-1999; 99US-0122487.  
 PR  
 XX  
 XX (GEGT) GENSET.  
 PA  
 XX  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI  
 XX  
 XX WPI; 2000-500361/45.  
 DR  
 XX  
 XX N-PSDB; AAC01774.  
 DR  
 XX  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNA that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PT  
 XX  
 XX Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (3'UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 CC  
 XX Sequence 111 AA;  
 SO  
 Query Match 31.0%; Score 323; DB 21; Length 111;  
 Best Local Similarity 59.0%; Pred. No. 7.8e-27;  
 Matches 62; Conservative 19; Mismatches 20; Indels 4; Gaps 1;  
 Oy 11 STLVACVDEVEFTNOEVEKEFEGLFRTYDECVTFOLFKSFRVRINFSPKSAARARIE 70  
 Db 11 SSLIACVANSIDIFSESSETRAKESLFTYDKDITFOYFKSFRVRINSPNSAADRLO 70  
 Oy 71 LHETQFRGKKLKLTPAQVQTPETDGDKLHLAPQAPQFLISPPS 115  
 Db 71 LHKTEFLGKEMKLYFAQTLHIGSS---HLAPPNPDQFLISPPA 111  
 RESULT 14  
 AAB58452  
 ID AAB58452 standard; Protein; 142 AA.  
 XX  
 XX AAB58452;  
 AC  
 XX  
 XX 14-MAR-2001 (first entry)  
 DT  
 XX  
 XX Lung cancer associated polypeptide sequence SEQ ID 790.  
 DE  
 XX  
 XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KM cardioactive; immunomodulatory; muscular active; vulnery;  
 KM gastrointestinal; nephrotoxic; antiinfective; gynecological;  
 KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KM proliferative disorder; wound healing; infectious disease.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO20005180-A2.  
 PN  
 XX  
 XX 21-SEP-2000.  
 PD  
 XX  
 XX 08-MAR-2000; 2000WO-US05918.  
 PP  
 XX  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX (ROSE/) ROSEN C A.  
 PA  
 XX  
 XX Ruben SM;  
 PI  
 XX  
 XX WPI; 2000-587514/55.  
 DR  
 XX  
 XX N-PSDB; AAF18328.  
 DR  
 XX  
 XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 PT  
 XX  
 XX Claim 11; Page 1325-1326; 1425pp; English.  
 PS  
 XX  
 XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1655.79 Seconds  
(without alignments)  
3480.130 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFRSYNFSSLIUCVAND.....RPKPKIIQTRRPEYTPHLS 198

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DBV=xlh  
-Q=/cgr2\_1/USPTO.spool/US09782953/runat\_11122002\_114429\_17340/app.query.fasta\_1.1173  
-DB=GenEmbl -DFT=fastap -SUFFIX=rge -MINMATCH=0\_1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09782953 @CGN\_1\_1\_2844 @runat\_11122002\_114429\_17340 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1043	100.0	597	6	AX365315 Sequence
2	1043	100.0	597	10	AF237790 Mus muscu
3	1028	98.6	2198	10	BC013551 Mus muscu
4	1025	98.3	2141	10	AF263239 Mus muscu
5	995.5	95.4	626	10	AB075973 Rattus no
6	992.5	95.2	2216	10	CG060263
7	970.5	93.0	2331	6	AX365321 Sequence
8	970.5	93.0	2346	9	HS085267
9	957.5	91.8	2348	6	AX281651
10	920.5	88.3	597	6	AX365312 Sequence
11	920.5	88.3	597	10	AF237789 Mus muscu
12	912.5	87.5	621	10	AF282255 Mus muscu
13	912.5	87.5	2125	10	AF260717 Mus muscu
14	890	85.3	2224	10	AF263240 Mus muscu
15	875	83.9	2212	6	AX365324 Sequence
16	875	83.9	2227	9	HS085266
17	875	83.9	2289	9	BC002864
18	875	83.9	2407	9	AK092184
19	868	83.2	599	6	AX365311
20	850.5	81.5	2284	9	HS085265
21	847.5	81.3	2173	6	AX410694
22	847.5	81.3	2173	9	HS028833
23	844.5	81.0	2174	6	AR034241
24	844.5	81.0	2679	9	AK055845
25	839.5	80.5	798	9	AF400429
26	807	77.4	562	9	HS053821
27	634.5	60.8	3295	10	AB061525
28	632.5	60.6	594	6	AX365318
29	632.5	60.6	594	10	AF237791
30	632.5	60.6	594	10	AF237887
31	627.5	60.2	3227	10	AB061524
32	623	59.7	934	6	AX074351
33	623	59.7	3159	6	AX420425
34	623	59.7	3261	9	AK090990
35	615.5	59.0	3184	6	AX329596
36	615.5	59.0	3184	6	AX333015
37	615.5	59.0	3184	6	AX365327
38	615.5	59.0	3184	6	AX420436
39	615.5	59.0	3184	9	HUMZAKI4
40	610	58.5	3240	9	AY034085
41	610	58.5	3253	9	AY034086
42	594.5	57.0	776	10	AF237888
43	592.5	56.8	828	6	AX365330
44	592.5	56.8	828	6	AX420437
45	592.5	56.8	828	9	AF176116

ALIGNMENTS

RESULT 1

AX365315  
 LOCUS AX365315 597 bp DNA linear PAT 15-FEB-2002  
 DEFINITION Sequence 5 from Patent WO0204491.  
 ACCESSION AX365315  
 VERSION AX365315.1 GI:18697045  
 KEYWORDS  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Williams, S.R. and Rothermel, B.  
 Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)  
 Patent: WO 0204491-A 5 17-JAN-2002;  
 Board of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rothermel, Beverly (US)  
 Location/Qualifiers

FEATURES  
 source  
 1..597  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 1..597  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD23809.1"  
 /db\_xref="GI:18697045"  
 /translation="MDPRDPSYFNSSLIACVANDVSESETRAKFSLFRTYDKDTT  
 FOYFKSRVRINFSNPLSADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPNPD  
 KQPLISPPASPVGMKQVEDATPVINDLVAISKLGPGEKYEHLAATDPTPSVVAVH  
 CESDQENEESEERMKRPKPIIQTRRPEYPIHLS"

CDS  
 156 c 142 g 129 t

BASE COUNT 170 a 156 c 142 g 129 t

ORIGIN

Alignment Scores:  
 Pred. No.: 2,92e-96 Length: 597  
 Score: 1043.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-782-953-7 (1-198) x AX365315 (1-597)

Qy 1 MetAappheaRgaspPheSerTyRanPheSerSerLeuIleAlaCyValAlaAsnAp 20  
 Db 1 ATGATTTTAAAGGAGCTTAAAGTAAATTTTACCTCCGATGCTGTCGCAAAAGAT 60  
 Qy 21 AppValPheSerGluSerGluThrArgAlaIlyPheGluSerLeuPheArgThyTyrap 40  
 Db 61 GATGCTTCAGCGAAAGTAGACACAGGGCCAAATTGAATCCCTTCAGAAATATATGAC 120  
 Qy 41 LybAappThThPhagIntPhyPheLySerPheLyArgValArgIleAnPheSerAn 60  
 Db 121 AAGGACACCACTTCAGATTTTAAAGCTTCACACGTCGCGATTAACCTTCAGCAAC 180  
 Qy 61 ProLeuSerAlaAlaAspAlaArgLeuAArgLeuAlaIlyThGluPheLeuGlu 80  
 Db 181 CCTTATCTGCAAGCCGATCCAGGCTCGGCTGCAACAACGAGTTCCTGGGGAAAGAA 240  
 Qy 81 MetLyLeuTyPheAlaGlnThrLeuHsIleGlySerSerHileuAlaProProAn 100  
 Db 241 ATGAAGTGTATTTGCTCAGACTTACACATAGAAATTCACACCTGCTCGGCCCAAT 300  
 Qy 101 ProAppLyGlnPheLeuLeSerProProAlaSerProProValGlyTrpLySGlnVal 120  
 Db 301 CCCGAAACACAGTCTCTCATCTCCCTCCGCTCTCCCTCGGTGGCTGGAAACAGATA 360  
 Qy 121 GluAspAlaThProValIleAnTyRaspLeuLeuTyAlaIleSerTySleuGlyPro 140  
 Db 361 GAAGATGCCACCCCTCATTAATTCAGATCTTTATATGCCATCTCCCAAGCTGGGCCCA 420  
 Qy 141 GtLyGluLyTyGtGluLeuHsAlaAlaTrnAapProThProSerValValAlaHsVal 160

Db 421 GGAGAGAGATATGACTGCATGCAGCAGACCCCACTCCAGTGTGTCACAGTG 480  
 Qy 161 CysGluSeraspGlnGluAangGluGluGluGluMetGluArgMetLyAspPro 180  
 Db 481 TGTGAGAGTGCACCAAGAGATGAGAGAGAGAGAGAGATGAGAGATGAAGAGACC 540  
 Qy 181 LybProLyValIleGlnThrArgArgProGluTyThrProIleHsIleuSer 198  
 Db 541 AAGCCAAATATCATTCAGACACGAGACCGAGTACACACCGATTCACCTTAC 594

RESULT 2  
 AF237790 597 bp mRNA linear ROD 12-Apr-2000  
 LOCUS AF237790  
 DEFINITION Mus musculus myocyte-enriched calcineurin interactin protein 1  
 ACCESSION AF237790  
 VERSION AF237790.1 GI:7542528  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 597)  
 Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.  
 A protein encoded within the Down Syndrome Critical Region is enriched in striated muscles and inhibits calcineurin signaling  
 J. Biol. Chem. (2000) In press  
 2 (bases 1 to 597)  
 Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.  
 Direct Submission  
 Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA  
 Location/Qualifiers

FEATURES  
 source  
 1..597  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 1..597  
 /note="similar to the EST with GenBank Accession Number AA200964, similar to Homo sapiens DSCR1"

CDS  
 156 c 142 g 129 t

ORIGIN

Alignment Scores:  
 Pred. No.: 2,92e-96 Length: 597  
 Score: 1043.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-782-953-7 (1-198) x AF237790 (1-597)

Qy 1 MetAappheaRgaspPheSerTyRanPheSerSerLeuIleAlaCyValAlaAsnAp 20  
 Db 1 ATGATTTTAAAGGAGCTTAAAGTAAATTTTACCTCCGATGCTGTCGCAAAAGAT 60  
 Qy 21 AppValPheSerGluSerGluThrArgAlaIlyPheGluSerLeuPheArgThyTyrap 40  
 Db 61 GATGCTTCAGCGAAAGTAGACACAGGGCCAAATTGAATCCCTTCAGAAATATATGAC 120



VERSION	AF263239.1	GI:8102011
KEYWORDS	Mus musculus.	
SOURCE	Mus musculus.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.	
REFERENCE	1 (bases 1 to 2141)	
AUTHORS	Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE	Down syndrome candidate region.1 (Dscr1), one of three alternatively spliced exon 1 transcripts	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2141)	
AUTHORS	Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO, Avia, Cascellideteils Km. 2,7, L'Hospitaler de Llobregat, Barcelona 08907, Spain	
FEATURES	Location/Qualifiers	
SOURCE	1..2141	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/chromosome="16"	
CDS	31..627	
	/note="Down syndrome candidate region 1; Dscr1; alternatively spliced exon 1"	
	/codon_start=1	
	/product="calcineurin inhibitor"	
	/protein_id="AA072701.1"	
	/db_xref="GI:8102012"	
	/translation="MHFRDYSNFSLLACVANDVPSESETRAKFESLFRYYDKDTTFOYFSPKRVIRINFSNPISADARLRHTEBLGKEMLYFAOTLIHIOSSHAPENPKOPILSPASPVGVKQVEDAPVINYLLVLAISLGGEXKELHAARDTPSVVHHCESDQNEBEEBEMRMRKPKIKITGRPERVPHLS"	
BASE COUNT	505 a	527 c 558 g 548 t 3 others
ORIGIN		
Alignment Scores:		
Pred. No.:	9 08e-94	Length: 2141
Score:	1025.00	Matches: 195
Percent Similarity:	98.99%	Conservative: 1
Best Local Similarity:	98.48%	Mismatch: 2
Query Match:	98.27%	Indels: 0
DB:	10	Gaps: 0
US-09-782-953-7 (1-198) x AF263239 (1-2141)		
Q1	MetAspPheArGApPheSerTyRAaPheSeSeSerLeuIleAcyValAlaAaMaP 20	
DB	31 ATGCATTTTAGGAGCATACATACATTTTAGCTCCCTGATGCTGTGGCAACGAT 90	
Q2	21 AspValPheSeSeGluSerGluThrArGAlaLylarPheGluSerLeuPheArGThTyRAAP 40	
DB	91 GATGCTTCAGCAAAAGTAGACACAGGGCCAAATTTGAATCCCTTCAGACATATGAC 150	
Q3	41 LylarPheThrArPheGlnTyRPhelYSerPheLylarGAlaArgIleAaPheSeArAa 60	
DB	151 AAGGACACCCCTTCAGATATTTAAGAGCTTCAAACGTCGCGATTAACCTTCAGCAAC 210	
Q4	61 ProLeuSerAlaAlaAaPAlaArGLeuArGLeuIleLylarThrGluPheLugIlyLyl 80	
DB	211 CCTTATCTGCACCGCATGCGAGCTCGGCTGCAACACCGAGTTCCTGGGGAAGAA 270	
Q5	81 MetLylLeuTyRPhelAgIlnThrLeuIleIleGlySeSeSerIleLeuAlaProProAa 100	
DB	271 ATCAAGTTGATATTTGCTCAGACTTTACATATGAGAGTTCAACCTCGGCTCCGCAAT 330	
Q6	101 ProAspLylGlnPheLeuIleSerProProAlaSerProProValGlyTYRPLYSGINVal 120	
DB	331 CCGGACAAACAGTTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGAAACAAAGTA 390	
Q7	121 GluAspAlaThrProValIleAaNTyrAspLeuLeuTyRAlaIleSerIlybLeuGlyPro 140	
DB	391 GAAGATCCACCCCGCATTAATTTAGATCTTTTATATGCACTTCCTCAAGCTGGGGCCA 450	

Oy	141	Glycyltyrosylglycylleucylisovalinalthreonylprothionoserinevalylvalhisval	160
Dd	451	GGAGAGAAGTATGAACGTGATCGACGCAGACACCCTCCAGTGCTGTGTTCACGTC	510
Oy	161	Cysglutserargininuanngluugluugluugluuinmetglnuaygmetylgargpro	180
Dd	511	TGTAGAGGTGCCAAGAAATGAGAGCAAGAGGAAAGATGGAGAAATGAAAGACCCC	570
Oy	181	LysProlysilileleginturargargproglinutythrprollehisleusSer	198
Dd	571	AAGCCCAAAAATCATCCAGACACCGSAGACC GGAGTACACACCCATCCACTCAGC	624
RESULT 5			
LOCUS	ABO75973	626 bp	RNA linear ROD 27-JUL-2002
DEFINITION	Rattus norvegicus mcip 1 mRNA for myocyte-enriched calcineurin-interacting protein 1, complete cds.		
ACCESSION	ABO75973		
VERSION	ABO75973.1		
KEYWORDS	GI:21998843		
SOURCE	Rattus norvegicus cDNA to mRNA.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 Odashima,M., Nagata,K., Obata,K., Somura,F., Izawa,H., Miyazaki,T., Murata,Y., Seo,H., Yamada,Y. and Yokota,M. rat myocyte-enriched calcineurin interactin protein 1, splice variant 4 mRNA, complete cds unpublished 2 (bases 1 to 626) Obata,K., Nagata,K., Odashima,M., Somura,F., Yamada,Y. and Yokota,M. Direct Submission Submitted (11-DEC-2001) Koji Obata, Nagoya University, Department of Clinical Pathophysiology; 65 Tsutsumi-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan (E-mail:obata@suru.med.nagoya-u.ac.jp, Tel:81-52-744-2577, Fax:81-52-744-2977)		
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
FEATURES	Location/Qualifiers		
source	1..626 /organism="Rattus norvegicus" /db_xref="taxon:10116"		
gene	1..626 /gene="mcip 1"		
CDS	19..618 /gene="mcip 1" /codon_start=1 /product="myocyte-enriched calcineurin-interacting protein 1"		
	1 /protein_id="BAC06443.1" /db_xref="gi:21998844" /translation="MHRRDFYNNSSLIACYANGDVSESTRAKESLPTDYDKDIT FQYEFRVRIRINSPLASADARLRHKTFGLGEMKLTYPAQTUHISSSLAPNP KOLFLSPASPVGWKCVEDATPVINDLYAIKLTGEGERELAAADTPPSVVVHH CESDEEEEEEERKKRPKPXIIQRREPEYPIHLS"		
BASE COUNT	182 a 163 c 154 g 127 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.gle-91	Length:	626
Score:	995.50	Matches:	192
Percent Similarity:	96.98%	Conservative:	1
Best Local Similarity:	96.48%	Mismatches:	5
Query Match:	95.45%	Indels:	1
DB:	10 Gaps:		1
US-09-782-953-7 (1-198) x ABO75973 (1-626)			
Oy	1	Metaspheargaphesertryanphieserleullealeacyvalalaanaap	20
Dd	19	ATGCATTTTTAGGAGCATTAACAATTTTACCTCCCTGATTCCTGTGTGGCAAACGCT	78

```

QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
DB 79 GATGCTCTCAGGAAAGTGAGACCAAGGCCAAATTTGAATCCCTCTTCAGGACGTATGAC 138

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValAtgIleAsnPheSerAsn 60
DB 139 AAGGACATCACTCTCCAGTATTTAAGAGCTTCAGAGCGTGCCGATTAATCTTCAGCAAC 198

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 199 CCCCTGTCTCGCGCGAGCGCCAGGCTGCGGTGCTAAGACAGAGATTCCTCGGGAAGGAG 258

QY 81 MetLysLeuTyrPheAlaGluThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
DB 259 ATGAACCTGTACTTTCGCCAGACTTTACATAGGAAGTTTACACCTGGCTCGGCCCAAT 318

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
DB 319 CCAGACAAACAGTTTCTCATCTCTCCCTTGCCTCTCCACTGTGTGGTGGAAACAGTA 378

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
DB 379 GAAGACGTACCCCAAGTACATAAATACAGTCTTTATATGCTATCTCCAGCTGGGACCA 438

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
DB 439 GGAGAGAGTACGAGCTACGCGAGCGACAGACACCAGCCCGAGCGTGGTCCACGTG 498

QY 161 CysGluSerAspGln--GluAsnGluGluGluGluMetGluAtgMetLysArg 179
DB 499 TGTGAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 559

QY 180 ProlysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
DB 559 CCCAAGCCAAAATCATCCAGACTCGGAGGCGAGTACACACCCATCCACCTCAGC 615

RESULT 6
CGU60263 2216 bp mRNA linear ROD 03-SEP-1997
LOCUS Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.
DEFINITION U60263
ACCESSION U60263.1 GI:2351390
VERSION 1
KEYWORDS Cricetulus griseus
SOURCE Cricetulus griseus
ORGANISM Cricetulus griseus
REFERENCE 1 (bases 1 to 2216)
AUTHORS Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and Davies,K.J.
TITLE Hamster adapt78 mRNA is a Down syndrome critical region homologue
JOURNAL Arch. Biochem. Biophys. 342 (1), 6-12 (1997)
MEDLINE 97329095
PUBMED 9185608
REFERENCE 2 (bases 1 to 2216)
AUTHORS Crawford,D.R., Leahy,K.L. and Davies,K.J.A.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical College, 47 New Scotland Avenue, Albany, NY 12208, USA
FEATURES
source 1..2216
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/cell_type="HA-1"
1..2216
/genes="adapt78"
70..563
/genes="adapt78"
/notes="similar to human DSCR1; an oxidant- and calcium-inducible mRNA"

```

```

BASE COUNT 544 a 515 c 564 g 593 t
ORIGIN
Alignment Scores:
Pred. No.: 1.79e-90 Length: 2216
Score: 992.50 Matches: 191
Percent Similarity: 97.47% Conservative: 2
Best Local Similarity: 96.46% Mismatches: 4
Query Match: 95.16% Indels: 1
DB: 10 Gaps: 1
US-09-782-953-7 (1-198) x CGU60263 (1-2216)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
DB 70 ATGCATTTTAGGAGCTTTAACTACAATTTTAGCTCCCTGATTCGCTGTGTGGCAACCGGT 129

QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
DB 130 GATGCTCTCAGCGAAAGTGAAACCGAGGCCAAATTTGAATCCCTCTTCAGGACGTATGAC 189

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
DB 190 AAGGACATCACTTTCCAGTATTTTAAAGAGCTTCAAAAGCTGTCCGAATAAATTCAGCAAC 249

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 250 CCTTATTCGAGCTGACGCCGCTGCGAGCTGCATAAGACGAGTTCCTGGGGGAAGGAA 309

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
DB 310 ATGAAGCTGTACTTCTCAGAGCTTACACATAGGAAGCTCACACCTGGCTCGGCCCAAT 369

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
DB 370 CCAGACAGCAGTTCCTCATCTCTCCCGGCTCTCCACCGAGTTGGTTGGGAAGCAAGTA 429

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
DB 430 GAAGACGCTACCCAGTCAATAATACGATCTTTATATGCTATCTCCAGCTGGGGCCA 489

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
DB 490 GGTGAGAGTATGAGTGCATGCAGCAACAGACACCACTCCAGTGTGTGTAGTCCACGTG 549

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
DB 550 TGTGAAGAGCGCAAGAGAAT---GAAGAGGAGGAGAGATCGAGAGAATGAAGAGACCC 606

QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
DB 607 AAGCCAAAATTTATCCAGCGAGGAGGCGGAGTACAGCTATCCACCTCAGC 660

RESULT 7
AX365321 2331 bp DNA linear PAT 15-FEB-2002
LOCUS AX365321
DEFINITION Sequence 11 from Patent WO0204491.
ACCESSION AX365321
VERSION AX365321.1 GI:18697049
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

AUTHORS Williams, S.R. and Rothermel, B.  
TITLE Methods and compositions relating to muscle selective calcineurin  
interacting protein (mcip)  
JOURNAL Patent: WO 0204491-A 11 17-JAN-2002;  
Board of Regents, The University of Texas System (US) ; Williams,  
Sanders R. (US) ; Rothermel, Beverly (US)  
FEATURES Location/Qualifiers  
SOURCE 1..2331  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
144..737  
/note="unnamed protein product"  
CDS /codon\_start=1  
/protein\_id="CAD23811.1"  
/translation="MHFRNPNVSSSLIACVANSIDPSEETRAKESLPRTYDKDIT  
FOYRSPKRVKINSNPSADARQLKHTFKREKMLYPAOTLHGSSHLAPND  
KQFLISPPASPPVGMKVEDATPVINDLLVAISKGPGRKYEELHAATDTTPSVVHV  
CESDQKEEEMERMRPKRIIOTRRPEYPIHLIS"  
BASE COUNT 630 a 470 c 547 g 684 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,16e-88 Length: 2331  
Score: 970.50 Matches: 185  
Percent Similarity: 96.97% Conservative: 7  
Best Local Similarity: 93.43% Mismatches: 5  
Query Match: 93.05% Indels: 1  
DB: 6 Gaps: 1  
US-09-782-953-7 (1-198) x AX365321 (1-2331)  
QY 1 MetAapPhaArgAapPheSerTyraPheSerSerLeuIleAlaCyValAlaAenAap 20  
Db 144 ATGCATTATTAAGAACTTTAACTACAGTTTACCTCCCTGATTCCTGCGCAACACT 203  
QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 40  
Db 204 GATATCTTCGCGAAAGTGAACACAGGGCCAAATTGAGTCCCTCTTAGAGAGTATGAC 263  
QY 41 LysAapThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAenPheSerAap 60  
Db 264 AAGGACATCACCTTTCAGTATTTTAAAGCTTCMAACGAGTCAAGATTAACCTTCACGCAAC 323  
QY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisIleGlySerSerHisIleValAapProAen 100  
Db 324 CCTTCTCCGACAGATGCCAGGCTCCAGCTCCATTAAGCTAGTTTCTGCGAAAGGAA 383  
QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisIleValAapProAen 100  
Db 384 ATCAAGTTATATTTGGCTAGACCTTACATAGGAACTCACACCTGGCTCCGCCAAAT 443  
QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120  
Db 444 CCGAGACAAAGAGTTTTCGATCTCCCTCCGCTCCGCGAGGGATGAAACAAAGTG 503  
QY 121 GluAapAlaThrProValIleAenTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140  
Db 504 GAAGATCGAACCCAGTCATTAAGTATGATCTCTTAATATGCCATCTCCAGCTCGGGCCA 563  
QY 141 GlyGluLysTyrrGluLeuHisIleAlaAlaThrAapProThrProSerValValValHisVal 160  
Db 564 GGGGAAAGATGAATTCAGCAGCAGCAGTGCACCACTCCAGCGGTGGTCCATGTA 623  
QY 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
Db 624 TGTGAGAGTATCAAGAG---AAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGCT 680  
QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198  
Db 681 AAGCCAAATTAATTCAGACCAAGAGCCGAGAGTACACGCCGATCCACCTCAGC 734  
RESULT 8

HSU85267  
LOCUS 2346 bp mRNA linear PRI 19-APR-2000  
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,  
alternative exon 1, complete cds.  
ACCESSION U85267  
VERSION U85267.2 GI:7596915  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2346)  
AUTHORS Fuentes, J.J., Pritchard, M.A., Planas, A.M., Bosch, A., Petter, I. and  
Battivill, X.  
TITLE A new human gene from the Down syndrome critical region encodes a  
proline-rich protein highly expressed in fetal brain and heart  
JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)  
MEDLINE 96121593  
PUBMED 8595418  
REFERENCE 2 (bases 1 to 2346)  
AUTHORS Fuentes, J.J., Pritchard, M.A. and Battivill, X.  
TITLE Genomic organization, alternative splicing, and expression patterns  
of the DSCR1 (Down syndrome candidate region 1) gene  
JOURNAL Genomics 44 (3), 358-361 (1997)  
MEDLINE 97468152  
PUBMED 9325060  
REFERENCE 3 (bases 1 to 2346)  
AUTHORS Fuentes, J.J., Pritchard, M. and Battivill, X.  
TITLE Direct Submission  
JOURNAL Submitted (14-APR-1997) Genetica Molecular, Institut de Recerca  
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalc, Barcelona  
08907, Spain  
REFERENCE 4 (bases 1 to 2346)  
AUTHORS Fuentes, J.J., Pritchard, M. and Battivill, X.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca  
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalc, Barcelona  
08907, Spain  
REMARK Sequence update by submitter  
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612867.  
FEATURES  
SOURCE 1..2346  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1-q22.2"  
1..2346  
/gene="DSCR1"  
144..737  
/gene="DSCR1"  
/note="down syndrome candidate region 1; one of four  
alternatively spliced exon 1"  
CDS /codon\_start=1  
/protein\_id="AA884372.1"  
/db\_xref="GI:2612868"  
/translation="MHFRNPNVSSSLIACVANSIDPSEETRAKESLPRTYDKDIT  
FOYRSPKRVKINSNPSADARQLKHTFKREKMLYPAOTLHGSSHLAPND  
KQFLISPPASPPVGMKVEDATPVINDLLVAISKGPGRKYEELHAATDTTPSVVHV  
CESDQKEEEMERMRPKRIIOTRRPEYPIHLIS"  
BASE COUNT 645 a 470 c 547 g 684 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,16e-88 Length: 2346  
Score: 970.50 Matches: 185  
Percent Similarity: 96.97% Conservative: 7  
Best Local Similarity: 93.43% Mismatches: 5  
Query Match: 93.05% Indels: 1  
DB: 9 Gaps: 1  
US-09-782-953-7 (1-198) x HSU85267 (1-2346)  
QY 1 MetAapPhaArgAapPheSerTyraPheSerSerLeuIleAlaCyValAlaAenAap 20

```

Db 144 ATGCATTTAGAACTTAACTACAGTTTACCTTACCTGATGCTGTGGGCAACAGT 203
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 204 GATATCTTCAGCAAGAGTGAACACGAGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 263
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLleAsnPheSerAsn 60
Db 264 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACGAGTCAGAAATAAATCTCAGCAAC 323
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 324 CCTTCTCCGAGCAGATGCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 393
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLleGlySerSerHisLeuAlaProProAsn 100
Db 384 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 443
Qy 101 ProAspLysGlnPheLeuLleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 444 CCAGCAAGCAGTTTCGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAACAAGTG 503
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaLleSerLysLeuGlyPro 140
Db 504 GAAGATCGGACCCAGTCATAACTATGATCTCTATATGCCATCTCCAAGCTGGGCGCA 563
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 564 GGGGAAAGTATGAATTCAGCGAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 623
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 624 TGTGAGAGTATCAAGAG---AAGGAGGAAGAGAGAGAAATGGAAGAATGAGGAGACCT 680
Qy 181 LysProLysIlelleGlnThrArgArgProGluTyrThrProLleHisLeuSer 198
Db 681 AAGCCAAAATATTCAGACCGAGGAGCGGAGTACACCGCGATCCACCTCAGC 734

RESULT 9
AX281651
LOCUS
DEFINITION
Sequence 60 from Patent WO0177389.
ACCESSION
AX281651
VERSION
AX281651.1 GI:16608902
KEYWORDS
human.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE
Genes expressed in foam cell differentiation
JOURNAL
Patent: WO 0177389-A 60 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1. .2348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 042176.5"
BASE COUNT 638 a 473 c 550 g 687 t
ORIGIN

Alignment Scores:
Pred. No.: 6,52e-87 Length: 2348
Score: 957.50 Matches: 185
Percent Similarity: 96.48% Conservative: 7
Best Local Similarity: 92.96% Mismatches: 5
Query Match: 91.80% Indels: 2
DB: 6 Gaps: 1

US-09-782-953-7 (1-198) x AX281651 (1-2348)

```

```

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db 151 ATGCATTTAGAACTTAACTACAGTTTACCTTACCTGATGCTGTGGCAACAGT 210
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 211 GATATCTTCAGCAAGAGTGAACACGAGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 270
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLleAsnPheSerAsn 60
Db 271 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACGAGTCAGAAATAAATCTCAGCAAC 330
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 331 CCTTCTCCGAGCAGATGCAGGCTCCAGCTGCATAAAGACTGAGTTTCTGGGAAAGGAA 390
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLleGlySerSerHisLeuAlaProProAs 100
Db 391 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 450
Qy 100 nProAspLysGlnPheLeuLleSerProProAlaSerProProValGlyTyrLysGlnVa 120
Db 451 TCCAGACAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAACAAGT 510
Qy 120 lGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaLleSerLysLeuGlyPr 140
Db 511 GGAAGATCGACCCAGTCATAAATATGATCTCTTATATGCCATCTCCAAGCTGGGCGC 570
Qy 140 OGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVa 160
Db 571 AAGGGAAGAGTATGAATTCAGCGAGCTGACACCACTCCAGCGTGGTGGTCCATGT 630
Qy 160 lCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPr 180
Db 631 ATGTGAGAGTATCAAGAG---AAGGAGGAAGAGAGGAATGGAAGAATGAGGAGACC 697
Qy 180 oLysProLysIlelleGlnThrArgArgProGluTyrThrProLleHisLeuSer 198
Db 698 TAAGCCAAAATATTCAGACCGAGGAGCGGAGTACACCGCGATCCACCTCAGC 742

RESULT 10
AX365312
LOCUS
DEFINITION
Sequence 2 from Patent WO0204491.
ACCESSION
AX365312
VERSION
AX365312.1 GI:18697043
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1
AUTHORS
Williams,S.R. and Rothermel,B.
TITLE
Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
JOURNAL
Patent: WO 0204491-A 2 17-JAN-2002;
Board of Regents, The University of Texas System (US) ; Williams,
Sanders R. (US) ; Rothermel, Beverly (US)
FEATURES
Location/Qualifiers
source
1. .597
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS
1. .597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23808.1"
/db_xref="GI:18697044"
/translation="MEEVDLQDLPSTAIACHLDPRFVVDGLCRAKFESLFRDYDKDTT
FOYFKSPKRVINFSNPLSADARLRLHKTFLCKEMKLYFAOTLHGSSHLAPNPDP
KQFLISPPASPVGWKQVEDATPVINDLLYLSKLGKGEKXELHAATDPTPSVVVHVH
CESQNEEEEEEMERMKRPKPKIOTRRPPTPIHLS"
BASE COUNT 160 a 170 c 151 g 116 t

```



ORIGIN

Alignment Scores:

Pred. No.:	6,596-84	Length:	597
Score:	920.50	Matches:	178
Percent Similarity:	92.42%	Conservative:	5
Best Local Similarity:	89.90%	Mismatches:	12
Query Match:	88.26%	Indels:	3
DB:	6	Gaps:	1

US-09-782-953-7 (1-198) x AX365312 (1-597)

OY 1 MetAapPheArGaPheSerTyraPheSerLeuIleAlaCyValAlaAsnAp 20

DB 10 GTGAGTCTGACGACCTGCGG-----AGCGCCACCATGCGCTGCCACCTGACCGG 60

OY 21 AAPValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThTyraP 40

DB 61 CGCGTGTCTGACGACGCTGTCGCGGCAAAATTGAAATCCCTTTCCAKAACATATGAC 120

OY 41 LysAapThThPheGlnTyPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 121 AAGGACACCACTTCACAGATTTTAAGACTTCAACGTCGCGATTAACTTCAGCAAC 180

OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 181 CCTTATCTGACAGCCGATGCCAGCTGCGGCTGCACAAAGCAGATTCTCGGGGAAGAA 240

OY 81 MetLysLeuTyPheLysIleGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 241 ATGAAGTTGATTATTTGCTCAGACTTACATAGAAAGTTCCACCTGCTCGGCCCAAT 300

OY 101 ProAepLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120

DB 301 CCGGACAAACAGTCTCTCATCTCCCTCGGCTCTCTCCGTTGGCTGGAAACAAGTA 360

OY 121 GluAapAlaThProValIleAsnTyraPheLeuLeuTyraIleSerLysLeuGlyPro 140

DB 361 GAAGATGCCACCCCGCTCAAAATACGATCTTTATATGCCATCTCCAAAGCTGGGGCCA 420

OY 141 GlyLysLysTyPheLysLeuHisIleAlaIleThrAspProThrProSerArgValAlaHisVal 160

DB 421 GGAGAAATATATGAACTGATGACGACGACAGACCCACCTCCAGTGGTGGTCCACGAG 480

OY 161 CysGluSerAepGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 481 TGTGAGAGTACCAACAGAAATGAGAGAAAGAAAGAAAGATGAGATGAAGAACACC 540

OY 181 LysProLysIleIleGlnThrArgArgProGlyTyrThrProIleHisLeuSer 198

DB 541 AAGCCCAAAATATCTCAGACACGAGACCGAGATACACCGATCCACCTTAGC 594

RESULT 11

AF237789 597 bp mRNA linear ROD 12-APR-2000

LOCUS AF237789

DEFINITION Mus musculus myocyte-enriched calineurin interactin protein 1

VERSION AF237789.1 GI:7542525

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 597)

Willems, B.A., Vega, R.B., Yang, J., Wu, H., Basseel-Duby, R.S. and Williams, R.S.

TITLE A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling

JOURNAL J. Biol. Chem. (2000) In press

REFERENCE 2 (bases 1 to 597)

Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Basseel-Duby, R.S. and Williams, R.S.

TITLE Direct Submission

JOURNAL Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

FEATURES

source

location/Qualifiers

1..597

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="16"

1..597

/note="Similar to Homo sapiens DSCR1; MGIPI; similar to EST sequence of GenBank Accession Number AY34360"

/product="myocyte-enriched calcineurin interactin protein 1 splice variant 1"

/protein\_id="AA63485.1"

/db\_xref="GI:7542525"

/translation="MEVDLQDLSATLACHLDPVVDGLCRAPESLFRYDKDT FQYPSKRRIRINSPNLSAADRLRHKTPEFGKMKLYFAOTLHGSHLAPNPD KQFLSPASPVGMKVEDATPVINDLVALSKLPGSEKVELHAATDPTPSVVAH CESDQENEEBEEEMERMRKPKIIOTRREYPIHLS"

BASE COUNT 160 a 170 c 151 g 116 t

ORIGIN

Alignment Scores:

Pred. No.:	6,596-84	Length:	597
Score:	920.50	Matches:	178
Percent Similarity:	92.42%	Conservative:	5
Best Local Similarity:	89.90%	Mismatches:	12
Query Match:	88.26%	Indels:	3
DB:	10	Gaps:	1

US-09-782-953-7 (1-198) x AF237789 (1-597)

OY 1 MetAapPheArGaPheSerTyraPheSerLeuIleAlaCyValAlaAsnAp 20

DB 10 GTGAGTCTGACGACCTGCGG-----AGCGCCACCATGCGCTGCCACCTGACCGG 60

OY 21 AAPValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThTyraP 40

DB 61 CGCGTGTCTGACGACGCTGTCGCGGCAAAATTGAAATCCCTTTCCAKAACATATGAC 120

OY 41 LysAapThThPheGlnTyPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 121 AAGGACACCACTTCACAGATTTTAAGACTTCAACGTCGCGATTAACTTCAGCAAC 180

OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 181 CCTTATCTGACAGCCGATGCCAGCTGCGGCTGCACAAAGCAGATTCTCGGGGAAGAA 240

OY 81 MetLysLeuTyPheLysIleGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 241 ATGAAGTTGATTATTTGCTCAGACTTACATAGAAAGTTCCACCTGCTCGGCCCAAT 300

OY 101 ProAepLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120

DB 301 CCGGACAAACAGTCTCTCATCTCCCTCGGCTCTCTCCGTTGGCTGGAAACAAGTA 360

OY 121 GluAapAlaThProValIleAsnTyraPheLeuLeuTyraIleSerLysLeuGlyPro 140

DB 361 GAAGATGCCACCCCGCTCAAAATACGATCTTTATATGCCATCTCCAAAGCTGGGGCCA 420

OY 141 GlyLysLysTyPheLysLeuHisIleAlaIleThrAspProThrProSerArgValAlaHisVal 160

DB 421 GGAGAAATATGAACTGATGACGACGACAGACCCACCTCCAGTGGTGGTCCACGAG 480

OY 161 CysGluSerAepGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 481 TGTGAGAGTACCAACAGAAATGAGAGAAAGAAAGATGAGATGAAGAACACC 540

OY 181 LysProLysIleIleGlnThrArgArgProGlyTyrThrProIleHisLeuSer 198

DB 541 AAGCCCAAAATATCTCAGACACGAGACCGAGATACACCGATCCACCTTAGC 594



```

RESULT 12
AF282255
LOCUS
DEFINITION
Mus musculus Down syndrome candidate region 1 protein (Dscr1) mRNA,
complete cds.
ACCESSION
AF282255
VERSION
AF282255.1 GI:9652249
KEYWORDS
Mus musculus.
SOURCE
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 621)
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
The murine DSCR1-like (Down syndrome candidate region 1) gene
family: conserved synteny with the human orthologous genes
Gene 257 (2), 223-232 (2000)
JOURNAL
MEDLINE
20534792
PUBMED
11080588
REFERENCE
2 (bases 1 to 621)
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
Direct Submission
Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia
Generale, Universite di Bologna, Via Belmeloro, 8, Bologna, BO
40126, Italy
FEATURES
Location/Qualifiers
1..621
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="16"
/sex="male"
/tissue_type="brain"
1..621
/gene="Dscr1"
8..604
/gene="Dscr1"
/note="similar to the Homo sapiens Down syndrome candidate
region 1 protein"
/codon_start=1
/product="Down syndrome candidate region 1 protein"
/protein_id="AA91461.1"
/db_xref="GI:9652250"
/tranlation="MREVDLQIPSAIACHLDPVVDGLCRAPESLFRVYDKDTT
FOYFKSKVRINFSNPLSADARLRLKTEFLGKMKLYFAQTLHGSSHLAPNPD
KQFLIFSPAPVPGWKOVEDATPVINDLYIAISKLGPGKVELHAATDTPTSVVVHV
CESDQNEEEEEEMERKPKPKIIQTRRPEVTPHLS"
BASE COUNT 166 a 179 c 159 g 117 t
ORIGIN

Alignment Scores:
Pred. No.: 4.43e-83 Length: 621
Score: 912.50 Matches: 177
Percent Similarity: 91.92% Conservative: 5
Best Local Similarity: 89.39% Mismatches: 13
Query Match: 87.49% Indels: 3
DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x AF282255 (1-621)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20
Db 17 GTGGATCTGCAGGACCTGCCG-----AGCCACCATCCCTGCCACCTGGACCCG 67

Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 68 CGCGTGTTCTGGACGCGCTGTGCCGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 127

Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLeuAsnPheSerAsn 60
Db 128 AAGGACACCACTTCAGTATTTAAGAGCTTCAACAGTGTCCGAGATAAACTTCAGCAAC 187

```

```

Qy 61 ProLeuSerAlaAlaAspAlaArgLeuAtcLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 188 CCTTATCTGCAGCCGATGCAGCTCGGCTGCACAAGACCGAGTTCTTGGGGAAGGAA 247

Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100
Db 248 ATGAAGTTGTATTGCTCAGACTTTTACACATAGGAAGTTTCACACCTGGCTCCGCCAAT 307

Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 308 CCCGACAAACAGTTCTCATCTCCCTCCGGCCTCTCTCCCTGGCTGGAACAAGTA 367

Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 368 GAAGATGCCACCCCGCTCATATAATTACGATCTTTTATATGCCATCTCCAAAGCTGGGCGCA 427

Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160
Db 428 GSAGAGAAAGTATGAACCTGCATGCAGCGACAGACACCACTCCCACTGTGTGTGCCAGTG 487

Qy 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
Db 488 TGTGAGAGTGCACCAAGAGATGAGGAGGAGGAGGAGATGGAGAGATGAAGAGACCC 547

Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 548 AAGCCCAAAATCATCCAGACAGCGAGACCGAGTACACACCCATCCCTCAGC 601

RESULT 13
AF260717
LOCUS
DEFINITION
Mus musculus Down syndrome candidate region 1 (Dscr1) mRNA, complete
cds.
ACCESSION
AF260717
VERSION
AF260717.1 GI:7839596
KEYWORDS
Mus musculus.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2125)
Casas, C., Martinez, S., Pritchard, M.A., Fuentes, J.J., Nadal, M.,
Guimera, J., Arbones, M., Florez, J., Soriano, E., Estivill, X. and
Alcantara, S.
Dscr1, a novel endogenous inhibitor of calcineurin signaling, is
expressed in the primitive ventricle of the heart and during
neurogenesis
Mech. Dev. 101 (1-2), 289-292 (2001)
JOURNAL
MEDLINE
21152920
PUBMED
11231093
REFERENCE
2 (bases 1 to 2125)
Fuentes, J.J., Pritchard, M., Pucharcas, C. and Estivill, X.
Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1
Unpublished
JOURNAL
REFERENCE
3 (bases 1 to 2125)
Fuentes, J.J., Pritchard, M., Pucharcas, C. and Estivill, X.
Direct Submission
TITLE
Submitted (26-APR-2000) Centre de Genetica Medica i Molecular, IRO,
Avia. Castelfideles Km. 2.7, L'Hospitalet de Llobregat, Barcelona
08907, Spain
FEATURES
Location/Qualifiers
1..2125
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..2125
/gene="Dscr1"
15..611
/gene="Dscr1"
/note="calcineurin inhibitor"
/codon_start=1
/product="Down syndrome candidate region 1"

```

```

/protein_id="AAF70343.1"
/db_xref="GI:7839597"
/translation="MEVDLDDPSATLADPRVVDGLCRKPFSLPRYDKOTT
FOYRSPKRVKINFSNPASADARLRHKPFLEKEMLYEAQTLHGSHLAPNPD
KQYRSPSPVGMKQVEDATPVINDLVAISKLGGEYELHAATDTTPSVVHVH
CESDQNEEBEEMERMKRPKIIQTRRPEYTPHLS"
BASE COUNT 487 a 536 c 563 g 536 t
ORIGIN
Alignment Scores:
Pred. No.: 1,99e-82 Length: 2125
Score: 912.50 Matches: 177
Percent Similarity: 91.92% Conservative: 15
Best Local Similarity: 89.39% Mismatches: 13
Query Match: 87.49% Indels: 3
DB: Gaps: 1
US-09-782-953-7 (1-198) x AAF260717 (1-2125)
QY 1 MetAaPheAaGpPheSerTyAaPheSerSerLeuIleAlaCyValAlaAsnAaP 20
Db GGGAGCTCTGCAGACCTGCGG-----AGCCACCATGCGCTGCACCTGCAGCCG 74
QY 21 AaPvaI)PheSerGluSerGluThrArgAlaAlaPheGluSerLeuPheArgThyTrpAaP 40
Db CGCGTGTTCGTGACGCGCTGTCGCGCCAAATTTGAATCCCTTCAGAAATATAGAC 134
QY 41 LysAaSerThrThrPheGluThrPheLeuSerPheLeuValAlaGluLeuPheSerAaP 60
Db AAGGACACCAACCTTCAGATTTTAAAGCTTCAACGTCGCGATTAACCTTCAGAAC 194
QY 61 ProLeuSerAlaAlaAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI) 80
Db 195 CCCTTATCTCCACCGCATGTCGACGCTGCGCTGCACAAAGCCAGTTCTCGGGAAGAA 254
QY 81 MetLysLeuTyThrPheAlaGluThrLeuHisIleGlySerSerHisIleuLeuAaPvaI)AaP 100
Db 255 ATGAAGTTGATTTTGTCTCAGACTTTACATAGAAATTCACACCTGCTCGCCCAAT 314
QY 101 ProAaPvaI)GluPheLeuIleSerProPvaI)AaPvaI)GlyTrpLysGluVal 120
Db 315 CCGGACAAACAGTTCTTCATCTCCCTCCGCTTCCTCCGTTGGCTGGAAACAGAT 374
QY 121 GluAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI) 140
Db 375 GAAGATGCCACCCCGCATTAATATAGATCTTTATATGCATCTTCACCAAGCTGGGCGCA 434
QY 141 GlyGluLysTyThrLeuHisIleAlaIleThrAaPvaI)AaPvaI)AaPvaI)AaPvaI) 160
Db 435 GGAGAGAAATATGAACTGTCGACGACACCACTCCAGTGTGTGTCACGTG 494
QY 161 CysGluSerAaPvaI)GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
Db 495 TGTAGAGTGCACAAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
QY 181 LysAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaP 198
Db 555 AAGCCAAAATCATCCAGACGACGAGACCGGAGTACACACCCATCCACCTCAGC 608
RESULT 14
AF263240 2224 bp mRNA linear ROD 30-MAY-2000
LOCUS AF263240
DEFINITION Mus musculus calcineurin inhibitor mRNA, complete cds,
alternatively spliced.
ACCESSION AF263240
VERSION AF263240.1 GI:8102013
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2224)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Pneumonia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.

```

```

TITLE Down syndrome candidate region 1 (Descr1), one of three
alternatively spliced exon 1 transcripts
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2224)
AUTHORS Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.
TITLE Direct Submision
JOURNAL Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
Avia, Castelldefels Km. 2,7, L'Hospital de Llobregat, Barcelona
08907, Spain
FEATURES
Source Location/Qualifiers
1..2224
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
354..710
/feature="Down syndrome candidate region 1; Descr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72702.1"
/db_xref="GI:8102014"
/translation="MKLYFAQTLHGSHLAPNPDKQFLISPPASPVGMKQVEDAT
PVINDLVAISKLGGEYELHAATDTTPSVVHVHVESDQNEEBEEMERMKRPK
KIQTRRPEYTPHLS"
BASE COUNT 519 a 554 c 586 g 562 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 3.91e-80 Length: 2224
Score: 890.00 Matches: 169
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 85.33% Indels: 0
DB: Gaps: 0
US-09-782-953-7 (1-198) x AAF263240 (1-2224)
QY 29 ArgAlaLysPheGluSerLeuPheArgThyTrpAaPvaI)AaPvaI)AaPvaI)AaPvaI) 48
Db 198 AGGGCCAAATTTGAATCCCTTCAGACATATGACAGACACCACTTCAGATATTT 257
QY 49 LysSerPheLysAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaP 68
Db 258 AAGACTTCAACCGTGTCCGATTAACCTTCAGAACCCCTTATGCGCGGACG 317
QY 69 LeuArgLeuHisIleGlyThrGluPheLeuGlyLysGluLysLeuTyThrPheAlaGluThr 88
Db 318 CTGCGGCTGCACAAAGACGAGTTCTCGGGAAGAAATGAAGTTGTATTTGCTCAGACT 377
QY 89 LeuHisIleGlySerSerHisIleuLeuAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI) 108
Db 378 TTACACATAGAAATTCACACTGCTCCGCCCAATCCCGCAAAAGATCTTCATCTCC 437
QY 109 ProPvaI)AaPvaI)GlyTrpLysGluValGluAaPvaI)AaPvaI)AaPvaI)AaPvaI) 128
Db 438 CTTCCGCTCTCCCTCCGTTGGCTGAAACAAATGAAGATGCAACCCCGTCAATAT 497
QY 129 TyrAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaP 148
Db 498 TACGATCTTTATGATCCATCTCCAAAGCTGGGCGGACGAGAGAGATGAACTGCATGCA 557
QY 149 AATTPAaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaP 168
Db 558 GGACAGACACCACTCCAGTGTGTCGACGTTGTGAGATGACCAAGAGAAATAG 617
QY 169 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 188
Db 618 GAGGAAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677
QY 189 ArgPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaPvaI)AaP 198
Db 678 AGACCGAGATGACACCACTTCACCTCAGC 707

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1118.55 Seconds  
(without alignments)  
2866.843 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFDFSYNFSLLIACVAND.....RPKPKIQTARPEYTHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09782953/runat.11122002.114429.17353/app.query.fasta.1.1173  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 -CGN_1_1583 -runat.11122002.114429.17353 -NCPU=6 -ICPU=3  
-NO_XLPEXT -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : EST.\*

1:	em_estba.*
2:	em_esthum.*
3:	em_estin.*
4:	em_estmu.*
5:	em_estov.*
6:	em_estpl.*
7:	em_estro.*
8:	em_htc.*
9:	gb_est1.*
10:	gb_est2.*
11:	gb_htc.*
12:	gb_est3.*
13:	gb_est4.*
14:	gb_est5.*
15:	em_estfun.*
16:	em_estom.*
17:	gb_gsa.*
18:	em_gsa_hum.*
19:	em_gsa_inv.*
20:	em_gsa_pin.*
21:	em_gsa_vrt.*
22:	em_gsa_fun.*
23:	em_gsa_nam.*
24:	em_gsa_mus.*
25:	em_gsa_other.*
26:	em_gsa_pro.*
27:	em_gsa_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1021	97.9	1158	14	BQ895506	BQ895506 AGENCOURT
2	974	93.4	753	13	B1148584	B1148584 6029111995
3	970.5	93.0	837	9	AU124628	AU124628 AU124628
4	970.5	93.0	931	9	AL543576	AL543576 AU543576
5	970.5	93.0	939	9	AL546817	AL546817 AU546817
6	970.5	93.0	1041	13	BM450020	BM450020 AGENCOURT
7	963.5	92.4	946	14	BQ278576	BQ278576 AGENCOURT
8	961.5	92.2	718	9	AU131040	AU131040 AU131040
9	957.5	91.8	885	9	AL544755	AL544755 AU544755
10	941.5	90.3	662	13	BI464521	BI464521 603206117
11	938.5	90.0	939	9	AL554686	AL554686 AU554686
12	937	89.8	710	10	BB617325	BB617325 BB617325
13	934.5	89.6	906	9	AL544313	AL544313 AU544313
14	932.5	89.4	1078	13	BM541636	BM541636 AGENCOURT
15	929.5	89.1	740	13	BI463566	BI463566 603206210
16	925	88.7	950	12	BF782633	BF782633 602107462
17	924	88.6	665	13	BI219142	BI219142 602935788
18	909.5	87.2	811	12	BG570239	BG570239 602590641
19	906.5	86.9	769	13	BI767955	BI767955 603061320
20	904.5	86.7	827	13	BI102432	BI102432 602889619
21	898.5	86.1	980	12	BG296537	BG296537 602394441
22	896	85.9	711	13	BG915568	BG915568 602815815
23	887.5	85.1	710	12	BG475986	BG475986 602521111
24	885.5	84.9	1507	11	AK010696	AK010696 Mus muscu
25	875	83.9	828	9	AL536447	AL536447 AL536447
26	875	83.9	890	9	AL538796	AL538796 AL538796
27	868	83.2	599	9	AA200984	AA200984 mu12ell.r
28	862	82.6	867	14	BQ427531	BQ427531 AGENCOURT
29	861.5	82.6	751	9	AL576189	AL576189 AL576189
30	850.5	81.5	1014	9	AL559594	AL559594 AL559594
31	848.5	81.4	689	9	AL652662	AL652662 AL652662
32	847	81.2	937	12	BE795722	BE795722 601590730
33	846	81.1	956	9	AL556803	AL556803 AL556803
34	842.5	80.8	875	12	BG574693	BG574693 602596709
35	840.5	80.6	642	14	BQ520533	BQ520533 NISC AL05
36	833.5	79.9	898	14	BQ733563	BQ733563 AGENCOURT
37	829.5	79.5	1047	12	BG287042	BG287042 602382338
38	829	79.5	923	9	AL551657	AL551657 AL551657
39	823.5	79.0	664	13	BJ061844	BJ061844 BJ061844
40	815.5	78.2	627	13	BI859506	BI859506 603385358
41	808	77.5	520	14	BQ749142	BQ749142 UT-M-FD0
42	793.5	76.1	822	12	BE895008	BE895008 601433720
43	791.5	75.9	986	12	BG828534	BG828534 602752653
44	784.5	75.2	599	13	BJ064020	BJ064020 BJ064020
45	779.5	74.7	596	13	BJ058360	BJ058360 BJ058360

# ALIGNMENTS

RESULT 1  
BQ895506

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BQ895506 1158 bp mRNA linear EST 16-AUG-2002  
AGENCOURT 8762889 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6329174  
5', mRNA sequence.

BQ895506  
BQ895506.1 GI:22287520

EST.  
house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1158)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)



```

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
Db 328 ATGAAGTG-TATTTGCTCAGACTTTACATAGGAAGTTCCACCTGGCTCCGCCCAAT 386

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 387 CCGGCAACAGTTCCTCATCTCCCTCCGCCCTCTCTCCCTGGTGGTGGAAACAAGTA 446

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 447 GAAGATGCCACCCCGTCGTCATTAATACGATCTTTATATGTCATCTCCAGCTGGGCCCA 506

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 507 GGAGAGAAGTATGAATGATGTCATGCGCAGACACACCACTCCAGTGTGGTCCACGTCG 566

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 567 TGTGAGAGTGACCAAGAGAGATGAGAGAGAGAGAGAGAGATGGAGAGAATGAAGAGACC 626

QY 181 LysProLysIleIleGlnThrArgProGluTyrThrProIleHisLeuSer 198
Db 627 AAGCCCAAAATCATCCAGACACGAGAGAGAGAGAGATGACACACCACTCCACCTCAGC 680

RESULT 3
LOCUS AU124628 NT2RM4 Homo sapiens cDNA clone NT2RM4000307 5', mRNA
DEFINITION AU124628
ACCESSION AU124628
VERSION AU124628.1 GI:10949344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ota.T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
HRI human cDNA project (Ota.T., Wakamatsu,A., Ozawa,M., Yamamoto,J.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1..837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000307"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 241 a 201 c 207 g 186 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-98 Length: 837
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7

```

```

Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1
DB: 9 Gaps: 1
US-09-782-953-7 (1-198) x AU124628 (1-837)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAenAap 20
Db 118 ATGCATTTTAGAAACTTTAACTACAGCTTTAGCTCCCTGATTCCTGTGTGGCAACAGT 177

QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 178 GATATCTTCAGCGAAGTGAACACAGGCGCAATTTGAGTCCCTCTTTTAGGACGATATGAC 237

QY 41 LysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 238 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAACAGAGTCAGAATAAATTCAGCAAC 297

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 298 CCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTCTGGGAAGGAA 357

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
Db 358 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACTGGCTCCGCCCAAT 417

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 418 CCAGCAAGAGCAGTTCTGATCTCCCTCCGCTCCGCCAGTGGGATGGAAACAAGTG 477

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 478 GAAGATGCCAGCCAGTATTAAGTATATCTTTATATGCACTCTCCAGCTGGGGCCA 537

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 538 GGGGAAAAGTAGTATGAATTCATCGACGAGTACACCACTCCAGCGCTGGTGCATGTA 597

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 598 TGTGAGAGTGATCAAGAG---AAGGAGGAGAAGAGGAAATGGAAAGAAATGAGGAGACCT 654

QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 655 AAGCCCAAAATTTATCCAGACGAGGCGGAGTACACCGCGATCCACCTCAGC 708

RESULT 4
LOCUS AL543576 931 bp mRNA linear EST 16-FEB-2001
DEFINITION AL543576 LTI NFL006 PL2 Homo sapiens cDNA clone CS0DI006YK09 5
prime, mRNA sequence.
ACCESSION AL543576
VERSION AL543576.1 GI:12876055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI006YK09"
/tissue_type="placenta"
FEATURES
source

```

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,76e-98 Length: 931  
Score: 970.50 Matches: 185  
Percent Similarity: 96.97% Conservative: 7  
Best Local Similarity: 93.43% Mismatches: 5  
Query Match: 93.05% Indels: 1  
Gaps: 1

US-09-782-953-7 (1-198) x AL54617 (1-931)

QY 1 MetAapPheArGpAapPheSerTYrAnPheSerSerLeuIleAlaCyValAlaAsnAap 20  
DB 50 ATGCATTTTAGAAACTTTTACTACAGTTTACCTCCCTGATGCTGTGGCAACAGT 109  
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40  
DB 110 GATATCTTCAGCAAACTGAAACCAAGGCGCAATTGAGTCCCTCTTATGAGAGATAGAC 169  
QY 41 LysAapThrPhaGlnTYrPheLysSerPheLysArgValArgIleAsnPheSerAan 60  
DB 170 AAGGACATCCCTTCGATTTTAAAGACTTCAAGATCGAATAAATTCACCAAC 229  
QY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
DB 230 CCTCTTCCGACAGATGCGAGCTCCAGCTGCATAGAGCTGATTTCTGGGAAAGGAA 289  
QY 81 MetLysLeuTYrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAan 100  
DB 290 ATGAACTTAAATTTTGTCTGACACTTACACATGAGAAAGCTCACCTGCTGGCCAAAT 349  
QY 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTYrLysGlnVal 120  
DB 350 CCGACAAAGCACTTCTGATCTCCCTCCGCTCCGCGAGTGGATGGAACAAAGTG 409  
QY 121 GluAapAlaThrProValIleAsnTYrAapLeuLeuTYrAlaIleSerLysLeuGlyPro 140  
DB 410 GAAGATGCGAGCCAGTCATTAATGATGCTTTATATGCAATCTCCAAAGCTGGGGCCA 469  
QY 141 GlyGluLysTYrGluLeuHisAlaAlaThrAapProThrProSerValAlaHisVal 160  
DB 470 GGGGAAAGATGATGACGCGACGCGACCTACACCACTCCACGCGTGGTCCAGTGA 529  
QY 161 CysGluSerAapGlnGluAanGluGluGluGluMetGluArgMetLysArgPro 180  
DB 530 TGTGAGACTGATCAAGG--AAGAGAGAAAGAGGAAATGAAAGATGAGAGAGACT 586  
QY 181 LysPheLysIleIleGlnThrArgArgProGluTYrThrProIleHisLeuSer 198  
DB 587 AAGCCAAAATTAATCCAGACGAGGCGCAGAGTACAGCGGATCCACTGAGC 640

RESULT 5  
AL54617 AL54617 939 bp mRNA linear EST 16-FEB-2001  
LOCUS AL54617.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1029YJ21.5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL54617  
VERSION AL54617.1 GI:12879906  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
1 (bases 1 to 939)  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS  
Li, W.B., Gruber, C., Jessup, J., and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

1..939  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1029YJ21"  
/clone\_1ib="LTI.NFL006.PL2"  
/issue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 263 a 219 c 238 g 218 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,77e-98 Length: 939  
Score: 970.50 Matches: 185  
Percent Similarity: 96.97% Conservative: 7  
Best Local Similarity: 93.43% Mismatches: 5  
Query Match: 93.05% Indels: 1  
Gaps: 1

US-09-782-953-7 (1-198) x AL54617 (1-939)

QY 1 MetAapPheArGpAapPheSerTYrAnPheSerSerLeuIleAlaCyValAlaAsnAap 20  
DB 53 ATGCATTTTAGAAACTTTTACTACAGTTTACCTCCCTGATGCTGTGGCAACAGT 112  
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40  
DB 113 GATATCTTCAGCAAACTGAAACCAAGGCGCAATTGAGTCCCTCTTATGAGAGATAGAC 172  
QY 41 LysAapThrPhaGlnTYrPheLysSerPheLysArgValArgIleAsnPheSerAan 60  
DB 173 AAGGACATCACTTTCAGATTTTAAAGCTTCAACAGATCGAATAAATTCACCAAC 232  
QY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
DB 233 CCTTCTCCGACAGATGCGAGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 292  
QY 81 MetLysLeuTYrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAan 100  
DB 293 ATGAACTTAAATTTTGTCTGACACTTACACATGAGAAAGCTCACCTGCTGGCCAAAT 352  
QY 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTYrLysGlnVal 120  
DB 353 CCGACAAAGCACTTCTGATCTCCCTCCGCTCCGCGAGTGGATGGAACAAAGTG 412  
QY 121 GluAapAlaThrProValIleAsnTYrAapLeuLeuTYrAlaIleSerLysLeuGlyPro 140  
DB 413 GAAGATGCGAGCCAGTCATTAATGATGCTTTATATGCAATCTCCAAAGCTGGGGCCA 472  
QY 141 GlyGluLysTYrGluLeuHisAlaAlaThrAapProThrProSerValAlaHisVal 160  
DB 473 GGGGAAAGATGATGACCGAGCTGACACCACTCCACGCGTGGTCCAGTGA 532  
QY 161 CysGluSerAapGlnGluAanGluGluGluGluMetGluArgMetLysArgPro 180



```

Db 533 TGTGAGTGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGATGAGGAGACCT 589
Qy 181 LysProlylelleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 590 AAGCAAAAATATTCAGACAGGAGGCGGAGTACAGCGGATCCACCTCAGC 643

RESULT 6
BM450020
LOCUS BM450020
DEFINITION AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
5', mRNA sequence.
ACCESSION BM450020
VERSION BM450020.1 GI:18499060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL12204 row: d column: 10
High quality sequence stop: 643.
FEATURES
source
1..1041
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528025"
/clone_lib="NIH MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 282 a 248 c 274 g 236 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 28-98 Length: 1041
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7
Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1
DB: 13 Gaps: 1

US-09-782-953-7 (1-198) x BM450020 (1-1041)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db 84 ATGCATTTTAGAACTTTAACTACAGTTTAGCTCCTGATTCCTGTTGTTGGCAACAGT 143
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 144 GATATCTTCAGCAAGAGTGAACACAGGCGCAAAATTTGAGTCCTCTTTAGGACGTATGAC 203
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 204 AAGGACATCACCTTCAGTATTTTAGAGCTTCAACAGAGTCAAGATTAACATTCAGCAAC 263
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

```

```

Db 264 CCCTTCTCCGACGACAGATGCCAGGCTCCAGCTCCATAAGACTGAGTTTCTGGCAAGGAA 323
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 324 ATGAGATTATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCGGCCAAAT 383
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 384 CCAGACAAGCAGTTTCTGATCTCTCCCTCCGCTCTCCGACAGTGGATGGAACAAGTG 443
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 444 GAAGATGCCACCCCACTATAAATATGATCTCTTATATGCCATCTCCAGCTGGGGCCA 503
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 504 GGGGAAAAGTATGATTTGCACGAGCGAGTGCACACCTCCACGCGTGGTGGTCCATGTA 563
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 564 TGTGAGAGTGATCAACAG---AAGGAGGAAGAAGAGGAATGGAAGAATGAGGACCT 620
Qy 181 LysProLysIlelleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 621 AAGCCAAAATATTCAGACAGGAGGCGGAGTACACGCGATCCACCTCAGC 674

RESULT 7
BQ278576
LOCUS BQ278576
DEFINITION AGENCOURT_6773220 NIH_MGC_128 Homo sapiens cDNA clone IMAGE:5811711
5', mRNA sequence.
ACCESSION BQ278576
VERSION BQ278576.1 GI:20488772
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM2060 row: p column: 16
High quality sequence stop: 548.
FEATURES
source
1..946
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5811711"
/clone_lib="NIH MGC 128"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggcgcctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGAGTGGCCATTCAGCGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGACATG-3'. Full-length
enriched library was constructed using the Clontech

```

Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH MGC 126 and NIH MGC 127). Library created in the laboratory of T. Uedlin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC Library."

BASE COUNT 264 a 243 c 239 g 200 t

## ALIGNMENT SCORES:

Pred. No.:	1.09e-97	Length:	946
Score:	963.50	Matches:	184
Percent Similarity:	95.96%	Conservative:	6
Best Local Similarity:	92.93%	Mismatches:	7
Query Match:	92.38%	Indels:	1
DB:	14	Gaps:	1

US-09-782-953-7 (1-198) x BQ278576 (1-946)

```

OY 1 MetAapPhearGApPheSerTyraenPheSerSerleuIleAlaCyValAlaAap 20
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 76 ATGCATTTAGAAACTTTAACTTACAGTTTATGCTCCCTGATTGCTGTGGCAACAGT 135

OY 21 AapValPheSerGluSerGluThrAlaAlaValPheGluSerLeuPheArgThrTyraP 40
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 136 GATATCTTCAGCCAAAGTGAAACAGGCCAAATTGAGTCCCTTTAGAGAGTATGAC 195

OY 41 LysAapThrPheGlnTyrrPheLysSerPheLysArgValArgIleAapPheSerAen 60
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 196 AAGGACATCCCTTCAGTATTTTAAAGCTTCAACAGATCAAGATTAACCTTCAGCAAC 255

OY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 256 CCTTCTCCGACAGCATGCGCAGCTCCAGCTCATAGACTGTTCTGGGAAAGGAA 315

OY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 316 ATGAACTTATATTTGCTCGACCTTACACATAGGAAGCTCACACCTGCTCCGCAAAAT 375

OY 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnVal 120
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 376 CCGACACAGCAGTTTCTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 435

OY 121 GluAapAlaThrProValIleAenTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 436 GAAGATGCGACCCCGACATTAATGATCTCTTATATGCTCAAGCTCCGAGGAGCA 495

OY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAapProThrProSerValValHisVal 160
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 496 GGGGAAAGATATATGACGCGACGACGACGACGACGACGACGACGACGACGACGACGAC 555

OY 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 556 TGTGAGAGTATCAAGAAC---GAGGAGAAAGAGAAATGAGAAATGAGAGAGAGCT 612

OY 181 LysAapLysIleIleGlnThrArgAapProGlyTrpThrProIleHisLeuSer 198
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 613 AAGCCAAAATTAATCCGACGACGAGAGCCGAGTACAGCCGATCCACTCCAGC 666

RESULT 8
AUI31040 718 bp mRNA linear EST 01-ANG-2002
LOCUS AUI31040 NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA
DEFINITION sequence.
ACCESSION AUI31040
VERSION AUI31040.1 GI:10991394
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Oca.T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

```

TITLE  
JOURNAL  
COMMENT  
Isogai,T.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomic Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: Genomic@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix  
Research Institute; cDNA library construction; Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
Location/Qualifiers

FEATURES  
source  
1..718  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NT2RP3001895"  
/clone\_lib="NT2RP3"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/note="Vector: pME18SF3; mRNA from NT2 neuronal precursor  
cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 217 a 174 c 169 g 155 t 3 others

## ALIGNMENT SCORES:

Pred. No.:	1.33e-97	Length:	718
Score:	961.50	Matches:	183
Percent Similarity:	96.46%	Conservative:	8
Best Local Similarity:	92.42%	Mismatches:	6
Query Match:	92.19%	Indels:	1
DB:	9	Gaps:	1

US-09-782-953-7 (1-198) x AUI31040 (1-718)

```

OY 1 MetAapPhearGApPheSerTyrrAapPheSerSerleuIleAlaCyValAlaAap 20
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 102 ATGCATTTAGAAACTTTAACTTACAGTTTATGCTCCCTGATTGCTGTGGCAACAGT 161

OY 21 AapValPheSerGluSerGluThrAlaAlaValPheGluSerLeuPheArgThrTyraP 40
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 162 GATATCTTCAGCCAAAGTGAAACAGGCCAAATTGAGTCCCTTTAGAGAGTATGAC 221

OY 41 LysAapThrPheGlnTyrrPheLysSerPheLysArgValArgIleAapPheSerAen 60
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 222 AAGGACATCCCTTCAGTATTTTAAAGCTTCAACAGATCAAGATTAACCTTCAGCAAC 281

OY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 282 CCTTCTCCGACAGCATGCGCAGCTCCAGCTCATAGACTGATTTCTGGGAAAGGAA 341

OY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 342 ATGAACTTATATTTGCTCGACCTTACACATAGGAAGCTCACACTGCTCCGCAAAAT 401

OY 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnVal 120
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 402 CCAGACAGCAGTTTCTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 461

OY 121 GluAapAlaThrProValIleAenTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 462 GAAGATGCGACCCCGACATTAATGATCTCTTATATGCTCAAGCTCCGAGGAGCA 521

OY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAapProThrProSerValValHisVal 160
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 522 GGGGAAAGATATATGACGCGACGACGACGACGACGACGACGACGACGACGACGACGAC 581

OY 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
DB 582 TGTGAGAGTATCAAGAG---AAGGAGAAAGAGAAATGAGAAATGAGAGAGAGCT 638

```

Qy 181 LysProLysIlelleGlThrArgProGluTyrThrProIleHisLeuSer 198  
 Db 639 AAGCAAAATTTATCAGACAGGANGCGGAGTACACGCCGATCCACTTAAC 692

RESULT 9  
 AL544755  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 HOMO sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 885)  
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqraf@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1. 885  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1022YJ23"  
 /clone\_lib="LTI\_NFL006\_PL2"  
 /tissue\_type="placenta"  
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed by  
 Life Technologies. Contact : Feng Liang Life Technologies,  
 a division of Invitrogen 9800 Medical Center Drive  
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 247 a 211 c 228 g 198 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4.74e-97 Length: 885  
 Score: 957.50 Matches: 185  
 Percent Similarity: 96.48% Conservative: 7  
 Best Local Similarity: 92.96% Mismatches: 5  
 Query Match: 91.80% Indels: 2  
 DB: Gaps: 1

US-09-782-953-7 (1-198) x AL544755 (1-885)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerLeuIleAlaCysValAlaAsnAsp 20  
 Db 50 ATGCATTTTAAAGAACTTTAACTACAGTTTATGCTCCCTGATTCGCTGTGTGGCAACAGT 109

Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeupheArgThrTyrAsp 40  
 Db 110 GATATCTTCAGCGAAGAGTGAACACAGGCCAAATTTAGTGTCCTCTTTAGGACGTATGAC 169

Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 Db 170 AAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAAGATAAATTCAGCAAC 229

Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 Db 230 CCCTTCTCCGACGACAGTGCAGGCTCCAGCTGCATAAAGACTCAGTTTCTGGGAAAGGAA 289

Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
 Db 290 ATGAAGTTATATTTTGGCTCAGACCTTACACATAGGAAGCTCACACTGGCTCGCCAAT 349

Query Match: 90.27% Indels: 2  
DB: 13 Gaps: 1

US-09-782-953-7 (1-198) x B164521 (1-662)

```

OY 1 MetAapPheArGaPpHeSeTyRAnPheSeSeTyRAlaCyValAlaAsnAsp 20
DB 79 ATGCATTTAGAAACTTTAACTAACGATTTTACCTCCCTGATGCGTGTGCAACAGT 138
OY 21 AapValPheSeGluSeGluThrArgAlaValPheGluSerLeuPheArgThrTyR 40
DB 139 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGATCTCTTAAAGAGATAGAC 198
OY 41 LysAapThrThPheGlnTyRPhelYsSerPheLysArgValArgIleAsnPheSerAn 60
DB 199 AAGGACATCACCTTTACAGATTTTAAAGAGCTTCAACGAGTCAAGATTAACCTGCAAC 258
OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 259 CCTTCTCCGACAGATGCGAGCTCCAGCTCATAGACTGATGCTTGGGAAAGGAA 318
OY 81 MetLysLeuTyR-Phe-AlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAs 100
DB 319 ATGAAGTTATATTTGGCTCAGACCTTACACATAGAGAGCTCACCTGCTCCGCAAA 378
OY 100 nP-roApLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVa 120
DB 379 TCACAGACAGACATTTCTGATCTCTCCCTCCGCTCTCCGAGTGGATGGAACAGT 438
OY 120 LgluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyP 140
DB 439 GGAAGATGCGACCCAGCTCATTAATATGATCTTATATGCGATCTCCAGCTCGGGCC 498
OY 140 OGlyGluLysTyR-GluLeuHisLysAlaThrAspProThrProSerValValHisLysVa 160
DB 499 AGGGGAAATTAAGTAATTCAGACGACGAGTCAACACCTCCAGGTGGTGTGTCATGT 558
OY 160 LysGluSerAspGlnGluAsnGluGluGluGluGluMetGlnLysMetLysArgP 180
DB 559 ATGTGACAGTATCAAGG---AAGGAGGAAGAAAGAAATGAAAGATGAGAGAGACC 615
OY 180 OlyPProLysIleIleGlnThrArgArgProGluTyRThrProIle 195
DB 616 TAAGCCAAAATTATCCAGACGAGGAGGCGGAGTACACCCGATC 661

```

RESULT 11 939 bp mRNA linear EST 16-FEB-2001  
AL554686  
LOCUS AL554686 L17 NFLO06\_PU2 Homo sapiens cDNA clone CSOD10851F09 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL554686  
VERSION AL554686.1 GI:12895707  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 939)  
L1, W.B., Gruber, C., Jeejee, J. and Polyes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. .939  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSOD10851F09"  
/c1one\_11b="L17 NFLO06\_PU2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA

BASE COUNT 262 a 219 c 238 g 218 t 2 others

## ALIGNMENT SCORES:

Pred. No.: 6.84e-95 Length: 939  
Score: 938.50 Matches: 183  
Percent Similarity: 95.00% Conservative: 7  
Best local Similarity: 91.50% Mismatches: 7  
Query Match: 89.98% Indels: 3  
DB: 9 Gaps: 1

US-09-782-953-7 (1-198) x AL554686 (1-939)

```

OY 1 MetAapPheArGaPpHeSeTyRAnPheSeSeTyRAlaCyValAlaAsnAsp 20
DB 53 ATGCATTTAGAAACTTTAACTAACGATTTTACCTCCCTGATGCGTGTGCAACAGT 112
OY 21 AapValPheSeGluSeGluThrArg-AlaLysPheGluSerLeuPheArgThrTyR 40
DB 113 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGATCTCTTAAAGAGATAGAC 172
OY 40 sPlyAspThrThPheGlnTyRPhelYsSerPheLysArgValArgIleAsnPheSerA 60
DB 173 ACAGAGACATCACCTTACAGATTTTAAAGAGCTTCAACGAGTCAAGATTAACCTGCA 232
OY 60 nP-roLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
DB 223 ACCCTTCTCCGACAGATGCGAGCTCCAGCTCATAGAGAGTCAACCTGCTCGGCAAG 292
OY 80 LysLysLeuTyR-PheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProA 100
DB 293 AAATGAAGTTATATTTGCTCAGACCTTACACATAGAGAGTCAACCTGCTCGGCAAG 352
OY 100 nP-roApLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnV 120
DB 353 ATCCAGACAGAGATTTGATCTCTCCCTCCGCTCTCCGAGTGGATGGAACAG 412
OY 120 aLgluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyP 140
DB 413 TGGAAAGATCGACCCAGCTCATTAATATGATCTTATATGCGATCTCCAGAGCTGGGCG 472
OY 140 rOGlyGluLysTyR-GluLeuHisLysAlaThrAspProThrProSerValValHisLysV 160
DB 473 CAGGGGAAAGTATGAATTCAGACGAGTCAACACCTCCAGCTGGTGTGTCATGT 532
OY 160 aLysGluSerAspGlnGluAsnGluGluGluGluGluMetGlnLysMetLysArgP 180
DB 533 TATGTWAGGTATCAAGG---AAGGAGGAAGAAAGAAATGAAAGATGAGAGAGAC 589
OY 180 rOlyPProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLysSer 198
DB 590 CTAAAGCCAAAATTATCCAGACGAGGAGGCGGAGTACCCGATCCACCTCCAGAC 645

```

## RESULT 12

BB617325 710 bp mRNA linear EST 26-OCT-2001  
LOCUS BB617325 RIKEN full-length enriched, 11 days pregnant adult female  
DEFINITION ovary and uterus Mus musculus cDNA clone 5033405J71, mRNA sequence.  
ACCESSION BB617325  
VERSION BB617325.1 GI:16457186  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifestech.com URL :  
http://fulllength.invitrogen.com"



BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr. Web : www.genoscope.cns.fr.  
Location/Qualifiers

## FEATURES

SOURCE

1..906  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSOD1019YH04"  
/clone\_lib="HLTI\_NFL006\_Pl2"  
/tissue\_type="Placenta"  
/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 215 c 229 g 202 t 7 others

ORIGIN

## Alignment Scores:

Pred. No.: 1,84e-94 Length: 906  
Score: 934.50 Matches: 182  
Percent Similarity: 95.45% Conservative: 7  
Best Local Similarity: 91.92% Mismatches: 8  
Query Match: 89.60% Indels: 2  
Gaps: 1

US-09-782-953-7 (1-198) x AL544313 (1-906)

OY 1 MetAspPheArgAspPheSerTYrAsnPheSerSerLeuIleAlaCYaValAlaAsnAsp 20  
DB 45 ATGCATTTAGAACTTAACTTAAGCTTACAGTTTACCCCKATTTGCTGTGGCAACAGT 104  
OY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTYrAsp 40  
DB 105 GATATCTTCAGCGAAGTGAACAGGCGCAATTGAGTCCCTTTAGAGCATATGAC 164  
OY 41 LysAspThrThrPheGlnTYrPheLysSerPheLysValArgLysLeuPheSerAsn 60  
DB 165 AAGGACATCACTCTTCAGTATTTCAGCTTCAACAGAGTCAGAAATTAACCTTCAAGCAC 224  
OY 61 ProLeuSerIleAlaAspAlaArgLeuArgLeuIleLysThrGluPheLeuGlyLysGlu 80  
DB 225 CCTTCTCCGACAGATGCCAGCTCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 284  
OY 81 MetLysLeuTYrPheAlaGlnThrLeuIleIleGlySerSerHisLeuAlaProProAsn 100  
DB 285 ATGAAGTTATATTTTGGCTCAGACCTTACACATAGAAAGCTCACACCTGCGCCCAAT 344  
OY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTYrPlyGlnVal 120  
DB 345 CCAACAGAGCAAGTTCGATCTCCCTCCGCTCCGAGTGGAGTGAACAACTG 404  
OY 121 GluAspAlaThrProValIleAsnTYrAspLeuLeuTYrAlaIleSerIleLeuGlyPro 140  
DB 405 GAAAGTGCAGACCCAGTCATATACTATGATCTTATATGCTCATCAAGCTGGGCGCA 464  
OY 141 GlyLysLeuTYrGluLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160  
DB 465 GGGGAAAGTGAATTCACGCGCG-ACGACACACACTCCACGCTGGTGGTCCATGTA 523  
OY 161 CysGluSerArgGlnGluAsnGluGlnGluGluGluMetGluLysMetLysArgPro 180  
DB 524 TGTGAGTGTGATCAAGAG--AAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 580  
OY 181 LysProLysIleIleGlnThrArgArgProGluTYrThrProIleHisLeuSer 198  
DB 581 AAGCCAAAATTATCCAGACCAAGAGGCGGAGTACAGCCGATCCACTCAGC 634  
RESULT 14

BM541636 1078 bp mRNA linear EST 20-FEB-2002  
LOCUS  
AGENCOURT 6439234 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5531772  
DEFINITION  
5', mRNA Sequence.

## ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM541636  
BM541636.1 GI:18770450  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1078)  
NIH-MGC http://mgc.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.jnl.gov  
Plate: LHAM12213 row: p column: 13  
High quality sequence start: 256  
High quality sequence stop: 624.  
Location/Qualifiers

## FEATURES

SOURCE

1..1078  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5531772"  
/clone\_lib="NIH\_MGC\_71"  
/tissue\_type="telomysarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: oligo dT.  
Average insert size 2.1 kb.  
BASE COUNT 302 a 301 c 240 g 235 t

ORIGIN

## Alignment Scores:

Pred. No.: 3.78e-94 Length: 1078  
Score: 932.50 Matches: 183  
Percent Similarity: 95.50% Conservative: 8  
Best Local Similarity: 91.50% Mismatches: 6  
Query Match: 89.41% Indels: 3  
Gaps: 1

US-09-782-953-7 (1-198) x BM541636 (1-1078)

OY 1 MetAspPheArgAspPheSerTYrAsnPheSerSerLeuIleAlaCYaValAlaAsnAsp 20  
DB 64 ATGCATTTAGAACTTAACTTAAGCTTACAGTTTACCCCKATTTGCTGTGGCAACAGT 123  
OY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTYrAsp 40  
DB 124 GATATCTTCAGCGAAGTGAACAGGCGCAATTGAGTCCCTTTAGAGCATATGAC 183  
OY 41 LysAspThrThrPheGlnTYrPheLysSerPheLysValArgLysLeuPheSerAsn 60  
DB 184 AAGGACATCACTCTTCAGTATTTCAGCTTCAACAGAGTCAGAAATTAACCTTCAAGCAC 243  
OY 61 ProLeuSerIleAlaAspAlaArgLeuArgLeuIleLysThrGluPheLeuGlyLysGlu 80  
DB 244 CCTTCTCCGACAGATATCTAGAGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGA 303  
OY 80 MetLysLeuTYrPheAlaGlnThrLeuIleIleGlySerSerHisLeuAlaProProAsn 100  
DB 304 AATGAAGTTATATTTTGGCTCAGACCTTACACATAGAAAGCTTACAGCTGGTCCGCGCAA 363  
OY 100 nProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTYrLysGlnVal 120

```

Db 36% TCCAGACAAGCAGTTCTGATCTCCCTCCCGCCTCTCCGCCAGTGGGATGGAAACAAGT 423
Qy 120 lGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPr 140
Db 424 GGAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCC 483
Qy 140 oGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVa 160
Db 484 AGGGAAAGATGATGAATTGCAGCGAGGAGTGCACCACTCCCAAGCGTGGTGGTCCATGT 543
Qy 160 lCysGluSerAspGlnGluLeuGluGluGluGluMetGluArgMetLysArgPr 180
Db 544 ATGTGAGATGATCAAG---AAGGAGGAGAGAGAGAAATGGAAGAATGAGGAGCC 600
Qy 180 oLysProLysIleleGlnThrArgA-gProGluTyrThr-ProIleHisLeuSer 198
Db 601 TAAGCCAAAATTTATCCAGACCCAGGCGCGAGTACACGCCCGCATCCACTCAGC 656

RESULT 15
BI463566 740 bp mRNA linear EST 21-AUG-2001
LOCUS 603206210F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5271867 5',
DEFINITION mRNA sequence.
ACCESSION BI463566
VERSION BI463566.1 GI:15254222
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 740)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1686 row: c column: 04
High quality sequence stop: 736.

FEATURES
source
1..740
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5271867"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-xhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to R01 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
instituted by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 216 a 177 c 183 g 164 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 5,3e-94 Length: 740
Score: 929.50 Matches: 184
Percent Similarity: 95.95% Conservative: 6
Best Local Similarity: 92.93% Mismatches: 8
Query Match: 89.12% Indels: 2
Db: 13 Gaps: 0

```

Search completed: December 14, 2002, 22:28:11  
Job time.: 1121.55 secs

US-09-782-953-7 (1-198) x BI463566 (1-740)

```

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db 79 ATGCATTTTAGAAACCTTTAACTACAGTTTAGCTCCCTGATTGCGCTGTTGGCAACAGT 138
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 139 GATATCTTCAGCGAAGTGAACACCGGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 198
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 199 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCAGAAATAAATTCAGCAAC 258
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 259 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGAAAGGAA 318
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100
Db 319 ATGAAGTTATATTTTCTTCAGACCTTACACATAGGAAGCTCACACTGGCTCGGCCAAAT 378
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
Db 379 CCAGACAAGCAGTTTCTGATCTCCCTCTCCGCTCTCCGCGAGTGGATGGAACAAGTG 438
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 439 GAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCA 498
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 499 GGGAA-AAGTATGAATTGCAGCGACGAGTGCACACCACTCCAGCGTGGTGGTCCATGTA 557
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 558 TGTGAGGTGATCAAGAGAAGGAGGAGGAAGGAAAC--ATGGAAAGAATGAGGAGACT 615
Qy 181 LysProLysIleleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 616 AAGCCAAAATTTATCCAGACCCAGGAGCGGAGTACACCGCGATCCACTCAGC 669

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:40 ; Search time 13.0219 Seconds  
(without alignments)  
1461.736 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRPFSYNSLIACVAND.....RPRKIIQTRRPEYPIHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	355	34.0	239	T34305	hypothetical prote
2	140.5	13.5	315	T49641	nebula related pro
3	98.5	9.4	1274	D84485	probable retroel
4	95.5	9.2	163	S62444	conserved hypotet
5	94	9.0	318	C84651	hypothetical prote
6	92	8.8	992	T46337	hypothetical prote
7	87	8.3	1396	T10627	hypothetical prote
8	86.5	8.3	822	S30198	Na+/H+-exchanging
9	86.5	8.3	1020	B88414	hypothetical prote
10	85.5	8.2	561	JC2436	5'-nucleotidase (E
11	85	8.1	495	B96672	hypothetical prote
12	84.5	8.1	452	A25346	3', 5'-cyclic-nucle
13	84.5	8.1	719	S51739	transcription repr
14	83.5	8.0	797	A70453	glutamate ammonia
15	82.5	7.9	734	S37898	probable serine/th
16	82.5	7.9	748	AB1939	hypothetical prote
17	82	7.9	284	D48516	surfactant protein
18	82	7.9	370	D28211	larvicidal toxin 4
19	82	7.9	649	B81673	tail specific prot
20	82	7.9	697	E96752	hypothetical prote
21	82	7.9	1477	T13797	tumor suppressor pr
22	81.5	7.8	233	H86593	Yop translocation
23	81.5	7.8	233	D72030	type III secretion
24	81.5	7.8	528	C69323	conserved hypotet
25	81.5	7.8	802	T21315	hypothetical prote
26	81	7.8	151	T10768	latex allergen hev
27	81	7.8	1173	T31421	C-terminal domain
28	81	7.8	2237	T45115	N-type calcium cha
29	81	7.8	2339	A42566	omega-conotoxin-se

30 80.5 7.7 360 2 S12850 protein TPX - Ther  
31 80.5 7.7 401 2 T24929 hypothetical prote  
32 80.5 7.7 954 2 I61714 co-repressor prote  
33 80 7.7 370 2 A27098 larvicidal toxin 4  
34 80 7.7 455 2 B86427 hypothetical prote  
35 80 7.7 681 2 T17342 hypothetical prote  
36 79.5 7.6 559 2 A57474 extracellular matr  
37 79.5 7.6 587 2 JCS300 Ran GTPase activat  
38 79 7.6 324 2 T00518 hypothetical prote  
39 79 7.6 1216 2 T09224 spindle assembly c  
40 78.5 7.5 201 2 A36985 p-factor mating ph  
41 78.5 7.5 211 2 S37792 hypothetical prote  
42 78.5 7.5 290 2 T42526 hypothetical prote  
43 78.5 7.5 671 2 C96534 probable Poly-A B1  
44 78.5 7.5 1706 2 I84499 zinc finger protei  
45 78.5 7.5 1880 2 T18531 tractin - medicina

## ALIGNMENTS

### RESULT 1

T34305 hypothetical protein F54E7.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34305

R:Bentley, D.  
submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid F54E7.  
A:Reference number: Z21502  
A:Accession: T34305

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-239 <BEN>

A:Cross-references: EMBL:U00067; PIDN:AACT7519.1; GSPDB:GN00021; CESP:F54E7.7

A:Experimental source: strain Bristol N2; clone F54E7

C:Genetics:

A:Gene: CESP:F54E7.7

A:Map position: 3

A:Introns: 40/3; 78/2; 165/2; 205/1

Query Match 34.0%; Score 355; DB 2; Length 239;  
Best Local Similarity 40.0%; Pred. No. 8.5e-24;

Matches 72; Conservative 38; Mismatches 66; Indels 4; Gaps 3;

Qy 11 SSLIACVANDVFSESETRAKFSLRTYDKDTTFQVFKSKRVIRINFSNPLSAADRLR 70

Db 30 NAIIVTQVPEDVFNKODKANFSLFTQIEKDIHFDFLRSFRVRVIFSSPENATAKLI 89

Qy 71 LHKTEFLGKMKLYFAOTLHI--GSSHLAPPNPKOFLISPPASPPVGVKQVEDATPVI- 127

Db 90 VQGSFKGHELKAFQRIYMSANSQMLSPPLEKQFLISPPSPVGVGWEQTKDMPVVC 149

Qy 128 NYDLLYAIKSGPEKVELHAATDPTFVSVMVHVCSEQNEEEEEEMERMKRPRPKIIOT 187

Db 150 NFDLWALASPAIDKEVHGDELTAIVHPCEITID-VPSAIEMPRTPRPSSPSLQS 208

### RESULT 2

T49641

nebula related protein [imported] - Neurospora crassa

N:Alternate names: protein B5022.230

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49641

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49641

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <SCH>

A:Cross-references: EMBL:AL555932, GSPDB:GN00116; NCSP:BS022.230  
 A:Experimental source: BAC clone BS022; strain OR74A  
 C:Genetic:  
 A:Gene: NCSP:BS022.230  
 A:Map position: 6

Query Match 13.5%; Score 140.5; DB 2; Length 315;  
 Best Local Similarity 34.5%; Pred. No. 8.3e-057; Indels 19; Gaps 6;  
 Matches 48; Conservative 19; Mismatches 53;

OY 9 NFFSLIACANDVSESEFTRAKFESLFTYDKOTTFQY--FKSFKRVINFSNPLSAD 66  
 DB 69 NNSL-----DVSADNLQTRIRLISQT---APFAMSPKSPRRIIVTFPDEQAIA 118  
 OY 67 ARRLHKEFLGKEMKLYPAOTLIHGS---HLAPPNDKQFLISPPASPPVGMKO-VE 121  
 DB 119 VASVMDGAILGRCRCRVYGGPTPIDVSADKHLALPDAGKLFISPPSPPHDWEQRM 178  
 OY 122 DA--TPVINYDLVLAISK 138  
 DB 179 DAPNTVVAHEDLAELAKL 197

RESULT 3  
 DB4485  
 Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84485  
 R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;  
 eues, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84485  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1274 <STO>  
 A:Cross-references: GB:AE002093; NID:g4309763; PIDN:AAD15532.1; GSPDB:GN00139  
 C:Genetic:  
 A:Gene: AC2907400  
 A:Map position: 2

Query Match 9.4%; Score 98.5; DB 2; Length 1274;  
 Best Local Similarity 24.8%; Pred. No. 2.3; Indels 82; Gaps 9;  
 Matches 53; Conservative 28; Mismatches 51;

OY 25 ESETRAKFESLFTYDKOTTFQYKSFRRVINSNPLSADAR-LRLHKEFLGKE--M 81  
 DB 674 ESRREAFEGFRGSIENDV-----KQIKELKALADSKSSSYHDMYLAKTOTQODNP 728  
 OY 82 KLYVAQTLHT---GSSHLAPPND-----KQFLISPPA---SPPVGMKQ 119  
 DB 729 KVOQOQTPDVYKPKETNNQFATPSPPSKQADVGKKEKTELNLKPKGRGRKPSQPK 768  
 OY 120 VEDATPVNYDLVLAISKLGGEKYLHA---ATDPTSEVVVAVHVCESDQNEEEEM 174  
 DB 789 VSETPMKI-----TIKLLKSGQEKAEKEDSDVVDVTDVVEYEMMLPESDEDEET 842  
 OY 175 ERMKRPKPKIIQTR-----PEYTPHLIS 198  
 DB 843 ERLKSVKEIRLKYVLADPGRSLINFEAMPHTS 876

RESULT 4  
 S62444  
 conserved hypothetical protein SPAC136.15c - fission yeast (Schizosaccharomyces pombe)  
 N:Alternate names: hypothetical protein SPAC24B11.04c  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 10-Dec-1999  
 C:Accession: S62444; S62549; T37650; T38331  
 R:Odell, C.; Bowman, S.

submitted to the EMBL Data Library, October 1995  
 A:Reference number: S62430  
 A:Accession: S62444  
 A:Molecule type: DNA  
 A:Residues: 1-163 <OD>  
 A:Cross-references: EMBL:Z54308; NID:g1008985; PIDN:CAA91108.1; PID:g1009000  
 R:Odell, C.; Churcher, C.M.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: S62546  
 A:Accession: S62549

A:Molecule type: DNA  
 A:Residues: 1-163 <OD2>  
 A:Cross-references: EMBL:Z67757; NID:g1061288; PIDN:CAA91769.1; PID:g1061292  
 R:Odell, C.; Bowman, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z21734  
 A:Accession: T37650  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-163 <OD3>  
 A:Cross-references: EMBL:Z54308; PIDN:CAA91108.1; GSPDB:GN00066; SPDB:SPAC136.15c  
 R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z21786  
 A:Accession: T38331  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-163 <OD4>  
 A:Cross-references: EMBL:Z67757; PIDN:CAA91769.1; GSPDB:GN00066; SPDB:SPAC24B11.04c  
 A:Experimental source: strain 972h-; cosmid c24B11  
 C:Genetic:  
 A:Gene: SPAC24B11.04c; SPAC136.15c; SPAC24B11.04c  
 A:Map position: 1L  
 C:Superfamily: fission yeast hypothetical protein SPAC136.15c

Query Match 9.2%; Score 95.5; DB 2; Length 163;  
 Best Local Similarity 32.1%; Pred. No. 0.33; Indels 11; Gaps 4;  
 Matches 25; Conservative 17; Mismatches 25;

OY 96 LAPVPDKQFLISPPASPPVGMKQVEDATP--VINYLVAISKLGGEKYLHAATDP 152  
 DB 85 LQVPKFKGNWLISPPGSPVGMKQVEDATP--VINYLVAISKLGGEKYLHAATDP 140  
 OY 153 TPVVVAVHVCESDQNEEM 170  
 DB 141 GPQIVY----SEHNTKE 154

RESULT 5  
 C84651  
 hypothetical protein At2g25670 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: C84651  
 R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;  
 eues, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: C84651  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <STO>  
 A:Cross-references: GB:AE002093; NID:g4874305; PIDN:AAD31367.1; GSPDB:GN00139  
 C:Genetic:  
 A:Gene: At2g25670  
 A:Map position: 2

Query Match 9.0%; Score 94; DB 2; Length 318;  
 Best Local Similarity 25.2%; Pred. No. 1; Indels 30; Gaps 5;  
 Matches 38; Conservative 19; Mismatches 64;

QY 54 VRINFSNPLSADARLRLHKLHTEFLGKEMKLYFAOTLHGSSHLAPP--NPQKQFLISPPA 111  
 Db 13 ITIQSTNLFAALDTRKKKKKKKAGKSK-----GSSKSRPEPEPEPVVWATTP 62  
 QY 112 SPVGVKQVEDATPVINVDLLVAISKLGPGKVELHAATDPTFSPVVHVCSDDQF----- 166  
 Db 63 LKVKSWADIDDDDDDDYYATTAPQSGWSTSLPSH-----TDSKDVHVHSESEEDILD 117  
 QY 167 -----NNEEEEMERMKRPKPIIQTTRPE 191  
 Db 118 EGGDDVEEEOETEVQVHPEPEV--KKAPE 146  
 RESULT 6  
 T46337  
 Hypothetical protein DKFZp43402413.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C:Accession: T46337  
 R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, January 2000  
 A:Reference number: Z23037  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-992 <AAA>  
 A:Cross-references: EMBL:AL137265  
 A:Experimental source: adult testis; clone DKFZp43402413  
 C:Genetics:  
 A:Note: DKFZp43402413.1

Query Match 8.8%; Score 92; DB 2; Length 992;  
 Best Local Similarity 22.7%; Pred. No. 6.4;  
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;

QY 96 LAPPNPQKFLISPPAS-----PPVG-----WKQVEDATPVINVDLLVAIS----- 136  
 Db 245 LSPPLPHEERAQSPRSLSATEEPPQPEGQPEWKEAE-----LGEDSAASLSQLSLQR 300  
 QY 137 -----XKPG-----EKYELHAATDPTPSV----- 156  
 Db 301 EQAPSPPAACEKGKQHSQAEELGPGQGEAEEDPEKVAVSPTTPVSPVSRSTPVPAPPEQ 360  
 QY 157 -----VVHVCSQDQNEEEEMERMKRPKPIIQTTRPPTPIH 196  
 Db 361 LSEAAKXAEVAQVLEQDQRHLLSKQ-EKMQQLREKLQCEEEESILSLH 411

RESULT 7  
 T10627  
 Hypothetical protein T13K14.10 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
 C:Accession: T10627  
 R:Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke  
 submitted to the Protein Sequence Database, June 1999  
 A:Reference number: Z16991  
 A:Accession: T10627  
 A:Molecule type: DNA  
 A:Residues: 1-1396 <BEV>  
 A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.10  
 A:Experimental source: cultivar Columbia; BAC clone T13K14  
 C:Genetics:  
 A:Gene: ATSP:T13K14.10  
 A:Map position: 4  
 A:Introns: 145/1; 171/2; 196/1; 242/3; 269/3; 298/3; 347/3; 374/3; 414/3; 442/3; 505/2;  
 131/3; 1139/3; 1163/3; 1194/3; 1226/3; 1259/3; 1290/3; 1344/3

Query Match 8.3%; Score 87; DB 2; Length 1396;  
 Best Local Similarity 22.5%; Pred. No. 27;  
 Matches 49; Conservative 24; Mismatches 81; Indels 64; Gaps 7;

QY 18 ANDDFVSESETRAKFESLPRTYDKDTTFQYFKSKVRINFSNPLSADARLRLH----- 72  
 Db 1010 AMGDVYFESSKLPAGEYKLQYLURHENVELLEKQLQTVFIER--NMGEIRLNLHSEPDG 1067  
 QY 73 -----KTEFLGKEMKLYFAOTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDAT 124  
 Db 1068 PFTGNGAFKSVLMPGVKEAF-----YLGPPTKOKLPKNTPOGSMVVG----- 1110  
 QY 125 PVINVDLLVAISKLGPGKVELH-----AATDPTSPVVV----- 158  
 Db 1111 -EISYGGKLSFDEKGNPKNDPHRLVKLDAPEDKKAASAPTCKSVSRLEQEVRDTKI 1169  
 QY 159 -HVCESDQNEEEEMERMKRPKPIIQTTRPEYTP 195  
 Db 1170 KFLGNLKOETEERSEWRKL-----CTCLKSEYDPDTP 1203

# RESULT 8 S30198

Na+/H+-exchanging protein - Chinese hamster  
 N:Alternate names: Na+/H+ antiporter; Na+/H+ exchanger  
 C:Species: Cricetulus griseus (Chinese hamster)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 05-Nov-1999  
 C:Accession: S30198; S29541  
 R:Counillon, L.; Pouyssegur, J.  
 Biochim. Biophys. Acta 1172, 343-345, 1993  
 A:Title: Nucleotide sequence of the Chinese hamster Na(+)/H(+) exchanger NHE1.  
 A:Reference number: S30198; MUID:93192332; PMID:8383540  
 A:Accession: S30198  
 A:Molecule type: mRNA  
 A:Residues: 1-822 <COU>  
 A:Cross-references: EMBL:X68970; NID:G49472; PIDN:CAA48771.1; PID:G49473  
 C:Genetics:  
 A:Gene: NHE1

C:Keywords: glycoprotein; ion transport; membrane protein  
 F:76,374,414/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 8.3%; Score 86.5; DB 2; Length 822;  
 Best Local Similarity 23.7%; Pred. No. 15;  
 Matches 40; Conservative 26; Mismatches 64; Indels 39; Gaps 8;

QY 44 TQYFKSKVRINFSNPLSAA--DARLRLHTEFLGKEMKLYFAOTLH----- 90  
 Db 646 TRQRLSYNRHTL-VADPYEEAWNQMLLRQKARQLEQKMSNYLTVPAAKLDSPMRSAR 704  
 QY 91 IGSSHLA-PPNDKQFLISPPASPPVGMKQVEDATPVINVDLLVAISKLGPGKVELHAA 149  
 Db 705 IGSDFLAEPRKADLPVITIDPASP-----QSPESVDLVNEELKAKV-----LGVN 749  
 QY 150 TDPTSPVVHVCSQDQNEEEEMERMKRPKPII-----IOTRPEYTP 194  
 Db 750 RDPT-----RLTRGDEDEDEDEGVIMWRKEPSSPGTDVTPAPMYSF 793

# RESULT 9 B86414

Hypothetical protein F28N24.14 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: B86414  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 anser, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: B86414; MUID:21016719; PMID:11130712  
 A:Accession: B86414  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-1020 <STO>  
 A:Cross-references: GB:AE005172; NID:G9502422; PIDN:MAF68121.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 8.3%; Score 86.5; DB 2; Length 1020;  
 Best Local Similarity 21.3%; Pred. No. 20;  
 Matches 51; Conservative 35; Mismatches 76; Indels 77; Gaps 13;

21 DVFSESETRAKESLFRDYKDTTFQYKSPKRV-RINFSNPLSAD-----ARLRLHK 73  
 DB 747 DVLQASRLKLKKS-----DGNTYNTFSFQLLPETGSLPDSYSDDDTFCRSSPYMSD 800

74 TEPLG-----KEMLY--PAOTLHIGSSHLAPP----- 99  
 DB 801 DYLSDNHSLNSSEPWESDGHGREOLYDFHSRHVDNNAEHSPLGISESSCVAV 860

100 -----NPDQFLISPPASPPYGM-----KVEDATPVINYD--LLYAIISK--LGPGCE 142  
 DB 861 NUSYLNQPAEPL---PPPPPPQMMVSKTPSEKEMDKTOSLQGEALRFAFEKHISLPTA 917

143 KYELHA--ATDPTPVVAVVCSDDQNEEEEMERMKPKP---KIITRRPEYTP 195  
 DB 918 KQELPEWMTSAPKPEIKALT---KNVREBKOSANAKETETGDLQOIRTOQFNLRPV 972

RESULT 10  
 JC2436  
 5'-nucleotide (EC 3.1.3.5), cytosolic - human  
 N:Alternate names: cytosolic purine 5'-nucleotide  
 C:Species: Homo sapiens (man)  
 C>Date: 15-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 21-Jul-2000  
 C:Accession: JC2436  
 R:OKa, J.; Matsumoto, A.; Hosokawa, Y.; Inoue, S.  
 Biochem. Biophys. Res. Commun. 205, 917-922, 1994  
 A:Title: Molecular cloning of human cytosolic purine 5'-nucleotide.  
 A:Reference number: JC2436; MUID:95091838; PMID:7999131  
 A:Accession: JC2436  
 A:Molecule type: mRNA  
 A:Residues: 1-561 <OKA>  
 A:Cross-references: DDBJ:D38524; NID:G633070; PIDN:BA07529.1; PID:G633071  
 C:Comment: This enzyme is a soluble nucleotidase. It hydrolyzes IMP and other purine nucleotides. This enzyme has a critical role in the maintenance of a constant composition C:Genetics:  
 A:Gene: GDB:NT5CP  
 A:Cross-references: GDB:569367  
 C:Keywords: cytosol; phosphoric monoester hydrolase

Query Match 8.2%; Score 85.5; DB 2; Length 561;  
 Best Local Similarity 24.2%; Pred. No. 12;  
 Matches 47; Conservative 29; Mismatches 71; Indels 47; Gaps 10;

11 SGLIACVANDVF-----SSSETRAKESLFR-----TVDKDTTF-----QYKSPK 52  
 DB 386 SSLFEELQSDIFLAEYLKLLDSSSNRPDISIQRIRKVTYHDMDCYGMGSLFRSGS 445

53 RVRINFSNPISAADARLHKTEFLGKEMKLYFAOTLHIGSSHLAPPN-----DKQ 104  
 DB 446 KQTLFASQVMKPAD---LYAASFIN---LLYFPPSYLFRAAHYLWMEHSEVTEHTVDIN 498

105 FLISPPAS---PYGMQKVEDATPVINYDLVLAISKLGPGKVELHAATDPTSPVAVHC 161  
 DB 499 EMESPPLATFRNRTSVDFKD---TDYKRHQLTRISISEIKPMLFPL-----APOEITH-C 547

162 ESDQNEEEEME 175  
 DB 548 HDEDDDEEEEEE 561

RESULT 11  
 B96672  
 hypothetical protein F13011.18 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: B96672  
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso, Chih, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anen, N.R.; Hughes, B.; Huizer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marshall, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A66141; MUID:21016719; PMID:11130712  
 A:Accession: B96672  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <STO>  
 A:Cross-references: GB:AE005173; NID:G95042423; PIDN:AAD38262.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F13011.18  
 A:Map position: 1

Query Match 8.1%; Score 85; DB 2; Length 495;  
 Best Local Similarity 20.1%; Pred. No. 11;  
 Matches 35; Conservative 28; Mismatches 57; Indels 54; Gaps 5;

20 DVFSESETRAK--FESLFRDYKDTTFQYF-----KSFK---RVRINFSNPLSA 64  
 DB 162 EDINRKRQRHKLQTLTSETLDEAFRWMNKLKDFEEKGFKLRPYRVIGELMNLKV 221

65 ADARLRLHKTEFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPYGMQKVEDAT 124  
 DB 222 AEGRKFFVLQOKLRNAML-----VEMGEADPN 250

125 PVINYDLVLAISKLGPGKVELHAATDPTSPVAVVCSDDQNEEEEMERMK 178  
 DB 251 NPANTGVIERNVQGGEEERERLVAQGP-----KDDNDDDDDEEFDDMK 296

RESULT 12  
 A25346  
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) precursor - slime mold (Dictyost.  
 C:Species: Dictyostelium discoideum  
 C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 07-May-1999  
 C:Accession: A25373; A25346; A46548  
 R:Podgorsek, G.J.; Franke, J.; Faure, M.; Kessin, R.H.  
 Mol. Cell. Biol. 9, 3938-3950, 1989  
 A:Title: The cyclic nucleotide phosphodiesterase gene of Dictyostelium discoideum utilizes  
 A:Reference number: A25373; MUID:89384622; PMID:2779573  
 A:Accession: A25373  
 A:Molecule type: DNA  
 A:Residues: 1-452 <POD>  
 R:Jacombe, M.L.; Podgorsek, G.J.; Franke, J.; Kessin, R.H.  
 J. Biol. Chem. 261, 16811-16817, 1986  
 A:Title: Molecular cloning and developmental expression of the cyclic nucleotide phospho  
 A:Reference number: A25346; MUID:87057386; PMID:3023365  
 A:Accession: A25346  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <LAC>  
 R:Podgorsek, G.J.; Franke, J.; Kessin, R.H.  
 J. Gen. Microbiol. 132, 1043-1050, 1986  
 A:Title: Isolation of a cDNA encoding a portion of the cyclic nucleotide phosphodiester  
 A:Reference number: A46548; MUID:87010528; PMID:3020155  
 A:Accession: A46548  
 A:Molecule type: mRNA  
 A:Residues: 1-70, 'Y', '72 <PO2>  
 A:Cross-references: GB:M15738  
 C:Genetics:  
 A:Introns: 7/2  
 C:Keywords: phosphoric diester hydrolase  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-452/Product: 3',5'-cyclic-nucleotide phosphodiesterase #status predicted <MAT>

Query Match 8.1%; Score 84.5; DB 2; Length 452;  
Best Local Similarity 26.3%; Pred. No. 11;  
Matches 45; Conservative 19; Mismatches 66; Indels 41; Gaps 8;

QY 2 DPRDFSY-----NFSLIACVANDVFESESTRAKFSLFTYDKDTTFQY 47  
DB 80 DCDASYITPLCTGGDLDEGNLSSPELLTKXGSLNFIALDAGTVWGVRLL-----TFKY 135  
QY 48 FKSFKVRNFSNPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLI 107  
DB 136 FNTL-----FNITPSMAVLPEQR--TSWFLKHWMSYF-----IGHSL--DRVGGGLIV 182  
QY 108 SPASPPVGVKQVEDATPVINYDLLVAISKLGPEKVELHAATDPTSPVVV 158  
DB 183 SPEDYLAKNMIDVQ--PPINNGIMGLIRKLG-----FKPTDFTSSSIL 223

RESULT 13  
S51739  
transcription repressor ABEP1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 08-Oct-1999  
C:Accession: S60227; S51739  
R:He, G.P.; Muise, A.; Li, A.W.; Ro, H.S.  
Nature 378, 92-96, 1995  
A:Title: A eukaryotic transcriptional repressor with carboxypeptidase activity.  
A:Reference number: S60227; MUID:96061010; PMID:7477299  
A:Accession: S60227  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-719 <HE2>  
A:Cross-references: EMBL:X80478; NID:g607131; PIDN:CAA56648.1; PID:g607132

Query Match 8.1%; Score 84.5; DB 2; Length 719;  
Best Local Similarity 21.8%; Pred. No. 19;  
Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

QY 54 VRINFSNPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISP 110  
DB 585 LRVDPSRPMTQQORRQORLRLMRQML--RRLNSTAGPATSPTP-----ALMPP 637  
QY 111 ASP-----PVGWK--QVEDATPVINYDLLVAISKLGPEKVELHAATDPT 153  
DB 638 PSPTAITLRPWLPPTTTAGWSESETETTYEVVT-----EFETEYGTD-- 681

QY 154 PSVVHVHVCSDQNEEEEME 175  
DB 682 ----LEVERIEEEEEEMD 699

RESULT 14  
A70453  
glutamate ammonia ligase adenyl-yl-transferase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: A70453  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: A70453  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-797 <AQF>  
A:Cross-references: GB:AE000755; NID:g2984063; PIDN:AAC07604.1; PID:g2984064; GB:AE00065  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: glnE

Query Match 8.0%; Score 83.5; DB 2; Length 797;  
Best Local Similarity 26.6%; Pred. No. 27;

Matches 51; Conservative 25; Mismatches 63; Indels 53; Gaps 10;

QY 3 FRDFSYN--FSSLIACVANDVFESESTR-----AKFESLFRITYDK-----DITFOY 47  
DB 427 FTVFSVDYLHLVA--RNPDLVEDALTYRDPFKPEFKREKYRETINLSPENLYRR 484  
QY 48 FKSFKVRNFSNPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLI 107  
DB 485 FKTWNEVRIGLVLLKEEDRYEKLKK-----LYISLSTLA-----DFIL 523  
QY 108 SPASPPVGVKQVEDATPVINYDLLVAISKLGPEKVELHAATDPTSPVVHVHVCSDQEN 167  
DB 524 SNL-----WK-----DITLENTALLYALGKLG---SRELNFNSOLD---LVFAVKSLEEK 569

QY 168 EEEEEEMERMKR 179  
DB 569 EKVHEKAKELVR 580

RESULT 15  
S37998  
probable serine/threonine-specific protein kinase (EC 2.7.1.1) YKL168c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 24-Sep-1999  
C:Accession: S37998; S44593; S38413  
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.  
submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37976  
A:Accession: S37998  
A:Molecule type: DNA  
A:Residues: 1-734 <VAN>  
A:Cross-references: EMBL:Z28168; NID:g486294; PIDN:CAA82010.1; PID:g486295; MIPS:YKL168c  
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.  
Yeast 10, 25-33, 1994  
A:Title: Sequencing and analysis of a 20.5 kb DNA segment located on the left arm of yeast chromosome V.  
A:Reference number: S44593  
A:Accession: S44593  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-734 <VA2>  
A:Cross-references: EMBL:Z26878; NID:g407503; PIDN:CAA81519.1; PID:g407514  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:KK08  
A:Cross-references: SGD:S0001651; MIPS:YKL168c  
A:Map position: 11L  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:420-720/Domain: protein kinase homology <KIN>  
F:428-436/Region: protein kinase ATP-binding motif  
F:573/Active site: Asp #status predicted

Query Match 7.9%; Score 82.5; DB 2; Length 734;  
Best Local Similarity 21.4%; Pred. No. 30;  
Matches 42; Conservative 29; Mismatches 64; Indels 61; Gaps 8;

QY 11 SSLIACVANDVFESESTRAKFESLFTYDKDTT-----FOYFKSFKR 53  
DB 155 SSVSSCDSSNGTTSSSDSQWADSLDSDNDLTPYRGSKDKILKSKDRAPYNYIDDYK 214  
QY 54 VRI---NFSNPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISP 109  
DB 215 KALRRATSYNPLPS-----KQFYNERLYTRRS-----PDEESLESL 252  
QY 110 P--ASPPV-----GWKQVEDATPVINYDLLVAISKLGPEKVELH---AATDPTSPVV 157  
DB 253 PRFAGADVQCIIIEQNGFKVYEDSGSHEHNKLSGVIAKLEKGNLSLPAHQSGSLRPRLGIT 312

QY 158 V-----HVCSDDQEN 167  
DB 313 LSGLFKHKHNECDIEN 328

Sun Dec 15 08:38:43 2002

us-09-782-953-7.rpx

Search completed: December 11, 2002, 11:38:33  
Job time : 15.0219 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.7083 Seconds  
(without alignments)  
1651.161 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFDFSYNFSLLIACVAND.....RKPKLIQTRRPYTPHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	98.6	198	11 Q91WQ4	Q91WQ4 mus musculus
2	839.5	80.5	171	4 Q96R03	Q96R03 homo sapien
3	634.5	60.8	243	11 Q9VIP4	Q9VIP4 mus musculus
4	627.5	60.2	192	11 Q9VIP5	Q9VIP5 mus musculus
5	140.5	13.5	315	3 Q9P5S9	Q9P5S9 neurospora
6	107.5	10.3	249	3 Q9P4A1	Q9P4A1 cryptococcus
7	98.5	9.4	1274	10 Q9ZQK2	Q9ZQK2 arabidopsis
8	94	9.0	318	10 Q9SL96	Q9SL96 arabidopsis
9	94	9.0	318	10 Q9VZ43	Q9VZ43 arabidopsis
10	92.5	8.9	680	17 Q9ZY88	Q9ZY88 pyrobaculum
11	92	8.8	992	4 Q9NTH6	Q9NTH6 homo sapien
12	92	8.8	1455	4 Q9UPV0	Q9UPV0 homo sapien
13	90.5	8.7	4025	4 Q9NR13	Q9NR13 homo sapien
14	88.5	8.5	328	10 Q9LS03	Q9LS03 arabidopsis
15	88	8.4	828	4 Q9BQ15	Q9BQ15 homo sapien
16	87.5	8.4	327	5 Q9XZRI	Q9XZRI ciona intes

Q9VB52 drosophila  
Q99JH6 mus musculus  
Q9SUC7 arabidopsis  
Q9LIP4 arabidopsis  
Q9XIQ8 arabidopsis  
Q94C73 arabidopsis  
Q46411 bos taurus  
Q9D8G6 mus musculus  
Q61281 mus musculus  
Q9V3W3 drosophila  
Q67651 aquifex aeo  
Q8WCE9 drosophila  
Q9XZ25 anabaena sp  
Q9U5Y1 dictyosteli  
Q9PJV2 chlamydia m  
Q9CAI9 arabidopsis  
Q88442 mus musculus  
Q76931 drosophila  
Q92780 chlamydia p  
Q29668 archaeoglob  
Q19764 caenorhabdi  
Q9Y6N4 homo sapien  
Q83624 rattus norv  
Q9NR59 homo sapien  
Q9H7N4 homo sapien  
Q46122 drosophila  
Q62141 mus musculus  
Q9C752 arabidopsis  
Q9VHV8 drosophila

## ALIGNMENTS

### RESULT 1

Q91WQ4 ID Q91WQ4 PRELIMINARY; PRT; 198 AA.  
AC Q91WQ4;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Down syndrome critical region homolog 1 (human).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013551; AAH13551.1; -  
SQ SEQUENCE 198 AA; 22851 NW; F017C68F18ACC187 CRC64;

Query Match 98.6%; Score 1028; DB 11; Length 198;  
Best Local Similarity 99.0%; Pred. No. 6.9e-92;  
Matches 196; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDRFDFSYNFSLLIACVANDVFSESETRAKFESLFRITYDKDTTQYFKSFVRVNFNSN 60  
Db 1 MHRFDFSYNFSLLIACVANDVFSESETRAKFESLFRITYDKDTTQYFKSFVRVNFNSN 60  
Qy 61 PLSAADARLRHKTEFLGKMKLYFAOTLHIGSSHLAPNPKQFLISPPASPPVGWQV 120  
Db 61 PLSAADARLRHKTEFLGKMKLYFAOTLHIGSSHLAPNPKQFLISPPASPPVGWQV 120  
Qy 121 EDATPVINYDLYLAISKLGPKGVKVELHAATPTFSVVVHVCSDOENEEEMERMKRP 180  
Db 121 EDATPVINYDLYLAISKLGPKGVKVELHAATPTFSVVVHVCSDOENEEEMERMKRP 180  
Qy 181 KPKIQTTRRPYTPHLS 198  
Db 181 KPKIQTTRRPYTPHLS 198

## RESULT 2

Q96R03 PRELIMINARY; PRT; 171 AA.  
 AC Q96R03;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Down syndrome critical region protein 1.  
 GN DSCR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,  
 RA Qiang B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF400429; AAK92478.1;  
 SO SEQUENCE 171 AA; 19823 MW; 133001ABADP0BD9 CRC64;

Query Match 80.5%; Score 839.5; DB 4; Length 171;  
 Best Local Similarity 95.3%; Pred. No. 1.1e-73;  
 Matches 161; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 30 AFESLFRYDKDTTFQYKSPKRVINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 89  
 DB 4 AFESLFRYDKDTTFQYKSPKRVINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 63  
 QY 90 HIGSSHLAPNDKQFLISPPAPVGMKQVEDATPVINYLILYAIKLGPEKYLMAA 149  
 DB 64 HIGSSHLAPNDKQFLISPPAPVGMKQVEDATPVINYLILYAIKLGPEKYLMAA 123  
 QY 150 TDTPTSVVVHVCSDOENEEEMERMKRPKXITQTRPEYTPHLS 198  
 DB 124 TDTPTSVVVHVCSDOENEEEMERMKRPKXITQTRPEYTPHLS 171

## RESULT 3

Q8V1P4 PRELIMINARY; PRT; 243 AA.

AC Q8V1P4;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4 beta.  
 GN ZAKI-4 BETA.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,  
 RA "calcineurin inhibitory protein ZAKI-4."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB061525; BAB71956.1;  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; ttm; 1.  
 SO SEQUENCE 243 AA; 27332 MW; 2C08C1A810291851 CRC64;

Query Match 60.8%; Score 634.5; DB 11; Length 243;  
 Best Local Similarity 63.9%; Pred. No. 1.4e-53;  
 Matches 124; Conservative 22; Mismatches 39; Indels 9; Gaps 2;

QY 5 DFQYNSSLIACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSA 64  
 DB 51 DFQDPLPSLACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSA 110  
 QY 65 ADARLRHLKTEFLGKEMKLYFAQTLHIGSS----HLAPNDKQFLISPPAPVGMKQV 120

DB 111 ARARIELHETQFGKKLKYFAQVOTPETDGLHLAPPOPAKQFLISPPSPVGMKPI 170  
 QY 121 EDATPVINYLILYAIKLGPEKYLMAADTPPSVVHVCSDOENEEEMERMKRP 180  
 DB 171 SDATPVINYLILYAIKLGPEKYLMAADTPPSVVHVCSDOENEEEMERMKRP 225  
 QY 181 KPXITQTRPEYTP 194  
 DB 226 KPXITQTRPEYTP 239

## RESULT 4

Q8V1P5 PRELIMINARY; PRT; 192 AA.

AC Q8V1P5;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4.  
 GN ZAKI-4.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,  
 RA "calcineurin inhibitory protein ZAKI-4."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB061524; BAB71955.1;  
 SO SEQUENCE 192 AA; 21540 MW; B2DB9819F4B36A8 CRC64;

Query Match 60.2%; Score 627.5; DB 11; Length 192;  
 Best Local Similarity 63.8%; Pred. No. 4.8e-53;  
 Matches 120; Conservative 24; Mismatches 35; Indels 9; Gaps 2;

QY 11 SSLACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARL 70  
 DB 6 SSLACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARL 65  
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPNDKQFLISPPAPVGMKQVEDATPV 126  
 DB 66 LKTEFLGKEMKLYFAQTLHIGSS----HLAPNDKQFLISPPAPVGMKQVEDATPV 125  
 QY 127 INYDILYAIKLGPEKYLMAADTPPSVVHVCSDOENEEEMERMKRPKXITQ 186  
 DB 126 LKYDILYAIKLGPEKYLMAADTPPSVVHVCSDOENEEEMERMKRPKXITQ 180  
 QY 187 TRRPEYTP 194  
 DB 181 TRRPEYTP 188

## RESULT 5

Q9PS59 PRELIMINARY; PRT; 315 AA.

AC Q9PS59;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE Nebula related protein.  
 GN B5022.230.  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariales; Sordariaceae; Neurospora.  
 CX NCB1\_TaxID=5141;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hohenfeld J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR [2]  
 SO SEQUENCE FROM N.A.



RA German/Neurospora genome project;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL35932; CAB91442.1; -  
 SQ SEQUENCE 315 AA; 34194 MW; 76792987653C3AB7 CRC64;

Query Match 13.5%; Score 140.5; DB 3; Length 315;  
 Best Local Similarity 34.5%; Pred. NO. 1.7e-05;  
 Matches 48; Conservative 19; Mismatches 53; Indels 19; Gaps 6;  
 QY 9 NFSSLIACVANDVFSESETRAKFESLRTYDKDTTFV--FKSPKRVIRNFSNPLSAD 66  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 67 ARLRLHKTFFLCKENKMLTPAQLTHGSS-----HLAPPNDKQFLISPPASPPVGMKQ-VE 121  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 119 VRSVWDGEGAILGERCVFVGQPTIDVSAADKHLALPAGKLFITSPSPSPFPHDWEQRME 178  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 122 DA--TPVINYDLLVYAIKSL 138  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 179 DAPNTWHAEDLAEALAKL 197  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118

RESULT 6  
 Q9P4A1 PRELIMINARY; PRT; 249 AA.  
 ID Q9P4A1  
 AC Q9P4A1  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Calciurexin-binding protein.  
 GN CBP1.  
 OS Cryptococcus neoformans var. neoformans.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
 OX NCBI\_TaxID=40410;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H99;  
 RX MEDLINE=20359261; PubMed=10899116;  
 RA Gorlach J., Fox D.S., Cutler N.S., Cox G.M., Perfect J.R., Heitman J.;  
 RT Identification and characterization of a highly conserved calcineurin  
 binding protein, CBP1/calciurexin, in Cryptococcus neoformans.  
 RL EMBO J. 19:3618-3629(2000).  
 DR EMBL; AF210799; AAF63734.1; -  
 SQ SEQUENCE 249 AA; 27293 MW; 78757183FABEF09 CRC64;

Query Match 10.3%; Score 107.5; DB 3; Length 249;  
 Best Local Similarity 28.7%; Pred. NO. 0.019;  
 Matches 35; Conservative 16; Mismatches 44; Indels 27; Gaps 5;  
 QY 49 KSPKRVIRNFSNPLSADAR-----LRL-----HKTEFLGEMKLYPAQTUHI----- 91  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 57 RGFORAIWFSEBEAENAKQGDWKLDPVVGEEKVDNEGKLDIELVRLHLPPTT 116  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 92 -----GSSHLAPPNDKQFLISPPASPPVGMKQVEDATP---VINYDLLVYAIKSL 141  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 117 LNPDPATTHLAPPPLPHNFLISPPGSPGEPGWEPAEAPNRIILPDLQALTEILNSG 176  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 142 EK 143  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 177 SK 178  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118

RESULT 7  
 Q9ZQK2 PRELIMINARY; PRT; 1274 AA.  
 ID Q9ZQK2  
 AC Q9ZQK2  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Putative retroelement pol polyprotein.  
 GN AT2G07400.  
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Macon T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.B.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC06217; RAD15532.1; -  
 DR InterPro; IPR004312; ATHILA.  
 DR InterPro; IPR003653; SUMO\_protease.  
 DR Pfam; PF03078; ATHILA; 1.  
 DR Pfam; PF02902; Peptidase\_C48; 1.  
 KW Polyprotein.  
 SQ SEQUENCE 1274 AA; 145232 MW; A234285FDCB45631 CRC64;  
 Query Match 9.4%; Score 98.5; DB 10; Length 1274;  
 Best Local Similarity 24.8%; Pred. NO. 1.3;  
 Matches 53; Conservative 28; Mismatches 82; Indels 51; Gaps 9;  
 QY 25 ESETRAKFESLRTYDKDTTFVFKSKRVIRNFSNPLSADAR-LRLHKTFFLGKES--M 81  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 674 ESRVEAKPEGRFGSIENDV-----KQIKELKAIADSKSSVIRDMYLAKTQFQQNDP 728  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 82 KLYFAQTUHI-----GSSHLAPPNDK-----KQFLISPPA-----SPVGMKQ 119  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 729 KVQQTQTPDVPKPTNNQFATSPSPSKQADVKGKTEKLNELIQKPPGRGRKPSQPKK 788  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 120 VEDATPVIYNDLLVYAIKSLGPGKVELHA-----ATDTPSVVHVHVCESDQENEESEEM 174  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 789 VSPTEPKI-----TIKKLKSQEAKEAEDSDVVDVTDKVVSEYNMLPESDSEET 842  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 175 ERMKRPKPIQTRR-----PEYTPIHLS 198  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 QY 843 ERLKSVKEIRLTKVLKPLDGRSLINPEAMPHTTS 876  
 Db |||||-----DVESADNLTQIRLISQT---AFIFAWSPKSPRRIVITFFDEQAIA 118  
 RESULT 8  
 Q9SL96 PRELIMINARY; PRT; 318 AA.  
 ID Q9SL96  
 AC Q9SL96  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE At2g25670 protein.  
 GN AT2G25670.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Macon T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

RA Cronin L.A., Shen M., Vanek S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.,  
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 402:761-768 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006053; AAD1367.1;  
 SQ SEQUENCE 318 AA; 35128 MW; 2F2E2CD269FFD004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;  
 Best Local Similarity 25.2%; Pred. No. 0.55;  
 Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

OY 54 VRINFSPNSAADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAP--NPDKQFLISPPA 111  
 DB 13 ITIGSTNLPAALDRKKKKKSDKAGSK-----GSSKSREPEKEPEQVWVWAPTP 62  
 OY 112 SPVGVKQVEDATPVINVDLYAISKLGPEKYLHAATDPSPVVVVCSDOE----- 166  
 DB 63 LKYSWADIDDEDDDDYVATAPQSGWSTSLPSH-----TDSKDVHVESESEEDILD 117  
 OY 167 -----NEEEEMERMKRPKPIOTRRPE 191  
 DB 118 EGGDDVEEEOETEVOVHPEPV--KKAPE 146

RESULT 9  
 ID Q8V243 PRELIMINARY; PRT; 318 AA.

AC Q8V243;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 35.2 kDa protein.  
 GN A7G25670.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Tortum M., Wu H.C.,  
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,  
 RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,  
 RA Narusaka M., Nguyen A., Palm C.J., Sakurai T., Saitou M., Seki M.,  
 RA Shinn P., Southwick A., Shimozaki K., Davis R.W., Becker J.R.,  
 RA Theologis A.;  
 RT "Full length cDNA of gene At2g25670 (GI:15225169).";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY065267; AL38743.1;  
 KW Hypothetical protein.

DR EMBL; AY065267; AL38743.1;  
 SQ SEQUENCE 318 AA; 35154 MW; 2F2E2CC278FED004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;  
 Best Local Similarity 25.2%; Pred. No. 0.55;  
 Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

OY 54 VRINFSPNSAADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAP--NPDKQFLISPPA 111  
 DB 13 ITIGSTNLPAALDRKKKKKSDKAGSK-----GSSKSREPEKEPEQVWVWAPTP 62  
 OY 112 SPVGVKQVEDATPVINVDLYAISKLGPEKYLHAATDPSPVVVVCSDOE----- 166  
 DB 63 LKYSWADIDDEDDDDYVATAPQSGWSTSLPSH-----TDSKDVHVESESEEDILD 117

OY 167 -----NEEEEMERMKRPKPIOTRRPE 191  
 DB 118 EGGDDVEEEOETEVOVHPEPV--KKAPE 146

RESULT 10  
 ID Q8ZY88 PRELIMINARY; PRT; 680 AA.

AC Q8ZY88;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE DNA replication licensing factor (mcm).  
 GN PAE0901.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller U.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).  
 DR EMBL; AE009792; AAL63108.1;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR001208; MCM.  
 DR Pfam; PF00493; MCM\_1.  
 DR ProDom; PD001041; MCM; 1.  
 DR SMART; SM00382; AAA\_1.  
 DR SMART; SM00350; MCM; 1.  
 DR PROSITE; PS50051; MCM\_2; 1.  
 KW Complete proteome.

SQ SEQUENCE 680 AA; 76173 MW; 79409981D729D151 CRC64;  
 Query Match 8.9%; Score 92.5; DB 17; Length 680;  
 Best Local Similarity 22.7%; Pred. No. 2.1;  
 Matches 57; Conservative 32; Mismatches 89; Indels 73; Gaps 12;

OY 1 MPRFSPNSFSLICVANDVDFSESETRAFESLFTYDQTFQYFKSPKRYINFSN 60  
 DB 41 VDFHILMFDSLADLVVERPKLVPEADKVRREIVEKDEETAKALKRFFPR--GS 97

OY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAP--NPDKQFLISPPA 111  
 DB 98 PLSTV---LRLKRSYIGRLIKIGIVTRQTPPKHFLYKALYRCTGCGYEIELQLERH 154

OY 96 LAPP-----NPDKQFLISPPSPVGVKQY-----EDATP-----VINDLVVAI 135  
 DB 155 VEPFAPKCPKCGASKFTLVTELISQYIDWKVIVQERPELDPGOLPNSVEVLLDLD--V 212

OY 136 SLGSGPEKYLHAATDP-----PSVV-----VHVCSDQNEEEEMERMKRPK 181  
 DB 213 DTVKREDITSLGVVDLTLSLKKRPPIVTSYIGVAV---DTMKKELVERI--TKDE 267

OY 182 PKIQ--TRRPE 191  
 DB 268 QKLEISRFPD 278

RESULT 11  
 ID Q9NTH6 PRELIMINARY; PRT; 992 AA.

AC Q9NTH6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE Hypothetical 112.6 kDa protein (Fragment).  
 GN DKFZP43402413.

OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=TESTIS;
RA	Kohrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; ALI37265; CAB70664.1; -
DR	Hypothetical protein.
FT	NON TER
FW	1
SQ	SEQUENCE 992 AA; 112628 MW; 039DF5B1E97EF02F CRC64;
 Query Match Best Local Similarity 8.8%; Score 92; DB 4; Length 992; Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps	
QY	96 LAPNPDKQLISPAS-----PPVG-----WKQVEDATPVINDLLYAIS-----136
Db	245 LSPPLPHEERAQSPPSLATEEPPOGPGQPEWKEAE-----LGEDSAASLSLQLSLQR 300
QY	137 -----KLGPQ---EKVELHAATDTPSV-----156
Db	301 EQAPSPAAACEKGEHQSAEEELCPQEBAEDPEEKVAVSPTPPVPSEVRSTEPVAPPEQ 360
QY	157 -----VVHCSQDNEEEEEMRMKPKPKIIOTRRDEYTPIH 196
Db	361 LSEAALKAMEEAQAQLEQDORHLLESQ-EKWQLREKLCQEBEEILRLH 411
 RESULT 12	
ID	Q9UPVO PRELIMINARY; PRT: 1455 AA.
AC	Q9UPVO;
DT	01-MAY-2000 (T-EMBLrel. 13, Created)
DT	01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT	01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE	KIAA1052 protein.
GN	KIAA1052.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN;
RX	MEDLINE=99397452; PubMed=10470851;
RA	Kikuno R., Nagase T., Ishikawa K., Hirosewa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT	"Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."
RL	DNA Res. 6:197-205(1999).
DR	EMBL; AB028975; BAA83004.1; -
DR	InterPro; IPR001202; WW_Rsp5_WWP.
DR	Pfam; PF00397; WW; 1.
DR	SMART; SM00456; WW; 1.
DR	PROSITE; PS50020; WW_DOMAIN 2; 1.
SQ	SEQUENCE 1455 AA; 163543 MW; 7F48093100C34819 CRC64;
 Query Match Best Local Similarity 8.8%; Score 92; DB 4; Length 1455; Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps	
QY	96 LAPNPDKQLISPAS-----PPVG-----WKQVEDATPVINDLLYAIS-----136
Db	474 LSPPLPHEERAQSPPSLATEEPPOGPGQPEWKEAE-----LGEDSAASLSLQLSLQR 529
QY	137 -----KLGPQ---EKVELHAATDTPSV-----156
Db	530 EQAPSPAAACEKGEHQSAEEELCPQEBAEDPEEKVAVSPTPPVPSEVRSTEPVAPPEQ 589

OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC Oecosteleae; Brassicales; Brassicaceae; Arabidopsie.  
 OK NCBI\_TaxID=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=20277480; PubMed=10819329;  
 RA Nakamura Y.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones.";  
 RL DNA Res. 7:131-135(2000).  
 DR EMBL; AB026649; BAB01084.1; -;  
 DR InterPro; IPR001776; Aerolysin.  
 DR PROSITE; PS00274; AEROLYSIN; UNKNOWN 1.  
 SQ SEQUENCE 328 AA; 37438 MW; A43A5FE1AD0634AA CRC64;  
 Query Match 8.5%; Score 88.5; DB 10; Length 328;  
 Best Local Similarity 21.8%; Pred. No. 2;  
 Matches 46; Conservative 23; Mismatches 71; Indels 71; Gaps 8;  
 QY 46 QYFKSKRVIN-FS-----NPLSADARLRLHKTFLGKEMLYFAQLHIGSSHL 96  
 DB 41 QYFSVDQDIRTNSFSFYGTHDPNP-SKVEQDRFDEDFCG-----FLAIGTLGT 90  
 QY 97 APPNPD-----KQFLISPPASPPVGMKVEDATPVINYDL 131  
 DB 91 DPETPKFSAMVAEDATGEIKEMAKLIAKKLDQFLKEYPEDTRSKRVKISNECPDQYDL 150  
 QY 132 -----LYAISKLGGEKY-ELHAATDPTPSV--VAVCESD 164  
 DB 151 FSSSIETKGSNGRVKKKSLTSLFKRQTVGGEPIEKHSTRDAIKRVFKLHGASSK 210  
 QY 165 QENEEEEEEMERMKRPKRIQTRPEYTPPI 195  
 DB 211 TRNDEDDSMKKKDKKXVCTCRKRVHPV 241  
 RESULT 15  
 Q9BQ15 PRELIMINARY: PRT; 828 AA.  
 AC Q9BQ15;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Hypothetical 89.1 kDa protein.  
 GN DKFZP761D221.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=AMYGDALA;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gaassenhuber J., Glassl S.,  
 RA Ansoege W., Boescher M., Bloecher H., Bauersachs S., Blum H.,  
 RA Lauber J., Duestehoeft A., Beyer A., Koehrer K., Strack N.,  
 RA Mewes H.W., Ottenwaeider B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Pouscka A.;  
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and  
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";  
 RL Genome Res. 11:422-435(2001).  
 DR EMBL; AL136561; CAB6496.1; -;  
 DR InterPro; IPR000087; Collagen.  
 DR InterPro; IPR002965; P\_rich\_extensn.

DR PRINTS; PR01217; PRICHEXTENS.  
 KW Hypothetical protein.  
 SQ SEQUENCE 828 AA; 89136 MW; E03E971EB7D1DC58 CRC64;  
 Query Match 8.4%; Score 88; DB 4; Length 828;  
 Best Local Similarity 20.7%; Pred. No. 7.5;  
 Matches 50; Conservative 21; Mismatches 79; Indels 92; Gaps 7;  
 QY 25 ESETRAFESLFRTYDKDTFOYFKSKRVIRINFSNPLSADARLRLHKTFLGKEMLY 84  
 DB 275 QSATEVKEIKLPSINDLSIFGPVLSKSAVN-----AEEKVHV 314  
 QY 85 FAQLHIGSSHLAPPNDKQFLSPASP-----PYG----- 116  
 DB 315 FSDT---SPEHVTPELPRREVSPATPDNPADSPAPGLGPGPGPGPPRNVL 371  
 QY 117 -----WKQVEDATPVINYDLLYAISK-----L 138  
 DB 372 SPLNLEEVQKVAEQT-FIKDYLETISSPKDGLGQRAATPPPPPTVRYTVSSPGGS 430  
 QY 139 GPGKEYELHAATDPTPSVVAVHVCESQENEEEEEEMERMKRP--KPKIQTREPEYTPPIH 196  
 DB 431 GPGPRTSGASSPARPATPLVPCRSITPPPPPPRPRPKLPKPGVGVSRPSPPIH 490  
 QY 197 LS 198  
 DB 491 SS 492

Search completed: December 11, 2002, 11:41:06  
 Job time : 28.7083 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.7184 Seconds  
(without alignments)  
858.887 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFDSYNFSSLIACVAND.....RPPKIIQTRRPTPIHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database 1:

A\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	198	23	AAE18912
2	970.5	93.0	197	23	AAE18914
3	920.5	88.3	198	23	AAE18911
4	875	83.9	197	23	AAE18915
5	835.5	80.1	171	20	AAW73898
6	632.5	60.6	197	23	AAE18913
7	623	59.7	234	22	AAE18918
8	623	59.7	255	23	AAW50760
9	615.5	59.0	192	23	AAE18916
10	592.5	56.8	241	23	AAE18917

11	591.5	56.7	212	23	AAE18918	Human MCIP associa
12	537	51.5	111	21	AAE18912	Human secreted pro
13	372	35.7	292	22	ABE18916	Drosophila melanog
14	369.5	35.4	142	21	AAE18915	Lung cancer associ
15	284	27.2	58	22	ABE18915	Peptide #2146 enco
16	284	27.2	58	22	ABE18915	Peptide #2178 enco
17	284	27.2	58	22	ABE18915	Peptide #4383 enco
18	284	27.2	58	22	ABE18915	Protein #2080 enco
19	284	27.2	58	22	ABE18915	Human brain expres
20	284	27.2	58	22	ABE18915	Human brain expres
21	284	27.2	58	22	ABE18915	Human bone marrow
22	284	27.2	58	22	ABE18915	Peptide #2105 enco
23	284	27.2	58	22	ABE18915	Peptide #4293 enco
24	284	27.2	58	22	ABE18915	Peptide #2211 enco
25	284	27.2	58	22	ABE18915	Peptide #4402 enco
26	284	27.2	58	22	ABE18915	Peptide #2088 enco
27	284	27.2	58	22	ABE18915	Peptide #4188 enco
28	284	27.2	58	22	ABE18915	Human peptide enco
29	284	27.2	58	22	ABE18915	Human peptide enco
30	271.5	26.0	56	22	ABE18915	Peptide #859 enco
31	271.5	26.0	56	22	ABE18915	Protein #812 enco
32	271.5	26.0	56	22	ABE18915	Human brain expres
33	271.5	26.0	56	22	ABE18915	Human bone marrow
34	271.5	26.0	56	22	ABE18915	Peptide #839 enco
35	271.5	26.0	56	22	ABE18915	Peptide #855 enco
36	271.5	26.0	56	22	ABE18915	Peptide #814 enco
37	271.5	26.0	56	22	ABE18915	Human peptide enco
38	271.5	26.0	56	22	ABE18915	Human MCIP1 leucin
39	98	9.4	19	23	AAE18921	Human secreted pro
40	94.5	9.1	915	22	ABE18923	Arabidopsis thalia
41	94	9.0	318	21	AAE18923	Arabidopsis thalia
42	94	9.0	318	21	AAE18923	Arabidopsis thalia
43	94	9.0	318	21	AAE18923	Arabidopsis thalia
44	94	9.0	318	21	AAE18923	Arabidopsis thalia
45	92	8.8	954	22	AAU14615	Novel bone marrow

#### ALIGNMENTS

RESULT 1  
AAE18912  
ID AAE18912 standard; Protein; 199 AA.  
XX AAE18912;  
AC  
DT 17-MAY-2002 (first entry)  
DE Mouse MCIP associated protein #3.  
XX  
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; mouse.  
XX  
OS Mus musculus.  
XX  
PN WO20020491-A2.  
XX  
PD 17-JAN-2002.  
XX  
PF 06-JUL-2001; 2001WO-US21662.  
XX  
PR 07-JUL-2000; 2000US-216601P.  
PR 13-FEB-2001; 2001US-0782953.  
XX  
PA (TEXA) UNIV TEXAS SYSTEM.  
PA (WILL) WILLIAMS S R.  
XX (ROTH) ROTHERMEL B.  
PI Williams SR, Rothermel B;  
XX WPI; 2002-179698/23.  
DR N-PSDB; AAD30153.  
XX

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 151-152; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated protein.  
 CC Note: This sequence has been described as human MCIP3 in the  
 CC specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 CC Sequence 198 AA;  
 SQ  
 Query Match 100.0%; Score 1043; DB 23; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 7,7e-108;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDRPDSYFSSLIACVANDVFSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60  
 DB 1 MDRPDSYFSSLIACVANDVFSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60  
 QY 61 PLSAADARLRLHKTFFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGMKOV 120  
 DB 61 PLSAADARLRLHKTFFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGMKOV 120  
 QY 121 EDATPVINYDLYAISKLGPEKYEELHAATDPTPSVVHVCSDOENEEBEEEMERMRP 180  
 DB 121 EDATPVINYDLYAISKLGPEKYEELHAATDPTPSVVHVCSDOENEEBEEEMERMRP 180  
 QY 181 KPKITQTRRPETPIHLS 198  
 DB 181 KPKITQTRRPETPIHLS 198

RESULT 2  
 AAE18914  
 ID AAE18914 standard; Protein; 197 AA.  
 XX  
 AC AAE18914;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated protein #1.  
 XX  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 XX 06-JUL-2001; 2001WO-US21662.  
 XX  
 XX 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.

XX  
 PI Williams SR, Rothermel B;  
 XX  
 XX WPI; 2002-17698/23.  
 DR N-PSDB; AAD30155.  
 DR  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Example 1; Page 157-158; 174pp; English.  
 CC  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as murine splice variant in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 CC Sequence 197 AA;  
 SQ  
 Query Match 93.0%; Score 970.5; DB 23; Length 197;  
 Best Local Similarity 93.4%; Pred. No. 9.3e-100;  
 Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRPDSYFSSLIACVANDVFSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60  
 DB 1 MDRPDSYFSSLIACVANDVFSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60  
 QY 61 PLSAADARLRLHKTFFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGMKOV 120  
 DB 61 PLSAADARLRLHKTFFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGMKOV 120  
 QY 121 EDATPVINYDLYAISKLGPEKYEELHAATDPTPSVVHVCSDOENEEBEEEMERMRP 180  
 DB 121 EDATPVINYDLYAISKLGPEKYEELHAATDPTPSVVHVCSDOENEEBEEEMERMRP 180  
 QY 181 KPKITQTRRPETPIHLS 198  
 DB 181 KPKITQTRRPETPIHLS 197

RESULT 3  
 AAE18911  
 ID AAE18911 standard; Protein; 198 AA.  
 XX  
 AC AAE18911;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated protein #2.  
 XX  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 XX 06-JUL-2001; 2001WO-US21662.  
 XX  
 XX (WILL/) WILLIAMS S R.  
 XX (ROTH/) ROTHERMEL B.



XX Homo sapiens.  
 OS US5869318-A.  
 XX  
 XX 09-FEB-1999.  
 PD  
 XX 07-JUN-1996; 96US-0665040.  
 PF  
 XX 07-JUN-1996; 95ES-0001140.  
 PR  
 XX (PALU/) PALLEJA X E.  
 PA  
 XX Puentes JI, Pallega XE, Pritchard M;  
 PI WPI; 1999-152781/13.  
 DR N-PSDB; AAX01282.  
 XX  
 PT DNA encoding foetal brain proteins - believed to be associated with  
 PT Down's syndrome  
 PS  
 XX Claim 4: Column 15-18; 19pp; English.  
 CC  
 CC This sequence is encoded by the Down's Syndrome critical region 1 (DSCR1)  
 CC gene of the invention. The DSCR1 gene was found to be located in the  
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient  
 CC expression of DSCR1 mRNA in the brains of young rats, compared to  
 CC expression levels in the brains of adult rats, suggests an important role  
 CC for DSCR1 during the development of the Central Nervous System (CNS), and  
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities  
 CC of mental retardation and/or heart defects as found in Down's syndrome  
 CC patients.  
 CC  
 SQ Sequence 171 AA;

Query Match 80.1%; Score 835.5; DB 20; Length 171;  
 Best Local Similarity 94.7%; Pred. No. 8.5e-85;  
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

OY 30 AKESLFRVYDKOTTFQYKSPKRVINFSNPASADARLAKHTFLGKEMKLYFAOTL 89  
 DB 4 AKESLFRVYDKOTTFQYKSPKRVINFSNPASADARLAKHTFLGKEMKLYFAOTL 63  
 OY 90 HISSSHLAPNPDKOFLISPPASPVGMKOVDPATVINYDLVATSKLGPGEYELHAA 149  
 DB 64 HISSSHLAPNPDKOFLISPPASPVGMKOVDPATVINYDLVATSKLGPGEYELHAA 123  
 OY 150 TDPTPSVVHVCSDDQENEEBEMERKPKPKIITQTRPEYTPHLS 198  
 DB 124 TDPTPSVVHVCSDDQENEEBEMERKPKPKIITQTRPEYTPHLS 171

RESULT 6  
 AAE18913  
 ID AAE18913 standard; Protein; 197 AA.  
 XX  
 AC AAE18913;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Mouse MCIP associated protein #4.  
 XX  
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; mouse.  
 XX  
 XX Mus musculus.  
 OS  
 XX  
 PN W0200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US921662.  
 XX

PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILLU/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 XX WPI; 2002-179698/23.  
 DR N-PSDB; AAD30154.  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 PS  
 XX Disclosure; Page 154; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated protein.  
 CC Note: This sequence has been described as mouse MCIP2 encoding DNA in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 CC  
 SQ Sequence 197 AA;

Query Match 60.6%; Score 632.5; DB 23; Length 197;  
 Best Local Similarity 64.4%; Pred. No. 4.5e-62;  
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

OY 11 SLLICVANDVDVFSSEIRAKFESLFRVYDKOTTFQYKSPKRVINFSNPASADARL 70  
 DB 11 SLLICVANDVDVFSSEIRAKFESLFRVYDKOTTFQYKSPKRVINFSNPASADARL 70  
 OY 71 LAKTEFLGKEMKLYFAOTLHIGSS---HLAPNPDKOFLISPPASPVGMKOVDPATV 126  
 DB 71 LAKTEFLGKEMKLYFAOTLHIGSS---HLAPNPDKOFLISPPASPVGMKOVDPATV 130  
 OY 127 INVDLLVATSKLGPGEYELHAAIDPTPSVYVHVCSDDQENEEBEMERKPKPKIITQ 186  
 DB 131 INVDLLVATSKLGPGEYELHAAIDPTPSVYVHVCSDDQENEEBEMERKPKPKIITQ 185  
 OY 187 TRRPEYTP 194  
 DB 186 TRRPEYTP 193

RESULT 7  
 AAB31788  
 ID AAB31788 standard; Protein; 234 AA.  
 XX  
 AC AAB31788;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Amino acid sequence of a human detoxification protein.  
 XX  
 KM Human; detoxification protein; DEXT; cancer; leukemia; melanoma;  
 KW adenocarcinoma; autoimmune disorder; inflammatory disorder;  
 KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
 KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
 KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis.





SQ Sequence 255 AA; 59.7%; Score 623; DB 23; Length 255;  
 Query Match Similarity 62.4%; Pred. No. 7,6e-61;  
 Best Local Similarity 62.4%; Pred. No. 7,6e-61;  
 Matches 123; Conservative 21; Mismatches 41; Indels 12; Gaps 3;

QY 2 DFRDSYNSPSSIIACVANDVSESETRAKESLFRITDKDTTFQYFKSPKVRINFGNP 61  
 DB 63 DFNDDP---NSLFPACVHOSVFEGBESKEKPEGLFRITDDCTFGLFSLFRVRINFGNP 119  
 QY 62 USAADARLRLHKEFLGKEMKLYFAQTLLHIGSS---HLAPPNPKQPLISPPASPPVGM 117  
 DB 120 KSAARARIEHLHETFRGKKLKYFAQVOTPERTDGDKLHLAPQPAKQFLISPPSSPPVGM 179  
 QY 118 KQVEDATPVINDLYALSKLGPGEKYEHLAATDPTPSVVHVCSDDQNEEBEEMERBM 177  
 DB 180 QGINATPVINDLYALVAKLGPGEKYEHLAAGTSTPSVVHVCSDDIEEDPK----- 234  
 QY 178 KRPKPKIOTRRPEYTP 194  
 DB 235 TSPKPKIOTRRPEYTP 251

RESULT 9  
 AAE18916  
 ID AAE18916 standard; Protein; 192 AA.  
 AC AAE18916;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #3.  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human.  
 XX Homo sapiens.  
 OS WO200204491-A2.  
 PN 17-JAN-2002.  
 PD 06-JUL-2001; 2001WO-US21662.  
 PF 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX Williams SR, Rothermel B;  
 PI MPI; 2002-179698/23.  
 DR N-PSDB; AAD30157.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX Example 1; Page 165-166; 174pp; English.

CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as human MCIP splice variant in  
 CC specification, however the sequence seems to be a MCIP associated  
 CC protein.

SQ Sequence 192 AA; 59.0%; Score 615.5; DB 23; Length 192;  
 Query Match Similarity 63.3%; Pred. No. 3,4e-60;  
 Best Local Similarity 63.3%; Pred. No. 3,4e-60;  
 Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

QY 11 SSLIACVANDVSESETRAKESLFRITDKDTTFQYFKSPKVRINFGNPISAAARLR 70  
 DB 6 STLAVACVDEVEFTNCEKKEKPEGLFRITDDCTFGLFSLFRVRINFGNPISAAARLR 65  
 QY 71 LHKTEFLGKEMKLYFAQTLLHIGSS---HLAPPNPKQPLISPPASPPVGMKQVEDATPV 126  
 DB 66 LHETQFRGKKLKYFAQVOTPERTDGDKLHLAPQPAKQFLISPPSSPPVGMQINATPV 125  
 QY 127 INVDLYALSKLGPGEKYEHLAATDPTPSVVHVCSDDQNEEBEEMERBMKRPKPKIIQ 186  
 DB 126 LNVDDLVAVAKLGPGEKYEHLAAGTSTPSVVHVCSDDIEEDPK-----TSPKPKIIQ 180  
 QY 187 TRRPEYTP 194  
 DB 181 TRRPEYTP 188

RESULT 10  
 AAE18917  
 ID AAE18917 standard; Protein; 241 AA.  
 AC AAE18917;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated protein #4.  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human.  
 XX Homo sapiens.  
 OS WO200204491-A2.  
 PN 17-JAN-2002.  
 PD 06-JUL-2001; 2001WO-US21662.  
 PF 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX Williams SR, Rothermel B;  
 PI MPI; 2002-179698/23.  
 DR N-PSDB; AAD30158.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX Example 1; Page 168-169; 174pp; English.

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC can be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the

```

transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterising the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is human MCIP associated protein. Note: This sequence has been described as splice variant of MCIP1 initiated by exon 4 in the specification, however the sequence seems to be a MCIP associated protein.
```

CC	SQ	Sequence	241 AA;
Query Match		56.8%; Score 592.5; DB 23; Length 241;	
Best Local Similarity		61.5%; Pred. No. 1.8e-57;	
Matches 120; Conservative		19; Mismatches 43; Indels 13; Gaps 3;	

  

QY	1	MDPRPSVNFSSLIACVANDVFSESETRAKFSLETTYDKOTTFFYFKSFKRVRFNSN	60
DB	40	MOLSDLP---TSLFACSVEHAFEAREQERFEALFTIYDDQVTQLFKSFRRVRINFSK	96
QY	61	PLSAADARLRHLKTEFLGEMKLYPAOTLHG----SSHLPAPNDPKOPLISPPASPPVG	116
DB	97	PEAARARIELHETDFNGOKLKLFPAQVMSEGVRDKSYLLPQPVKQLISPPASPPVG	156
QY	117	WKQVEDATPVINDYLIIAYISKLGPGKEYELHAATDTPSVVVHVCSDOENEEEEEEMER	176
DB	157	WQSESDAMPVINYDLLCAVS KLGFGEYELHAGTESTPSVVHVCSETESEE-----	210
QY	177	MKRPKKILOTTRPE 191	
DB	211	TKNPKOKIAQTRRPD 225	

  

RESULT 11  
AAE18918  
ID AAE18918 standard; Protein; 212 AA.  
XX AC AAE18918;  
XX XX  
DT 17-MAY-2002 (first entry)  
DE Human MCIP associated protein #5.  
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
XX heart failure; cardiomyopathy; heart disease; human.  
OS Homo sapiens.  
PN WO200204491-A2.  
XX XX  
XX 17-JAN-2002.  
XX 06-JUL-2001; 2001WO-US21662.  
PP PF  
PP 07-JUL-2000; 2000US-216601P.  
PPR 13-FEB-2001; 2001US-0782953.  
XX XX  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PPA (WILL/) WILLIAMS S R.  
PPA (ROTH/) ROTHERMEL B.  
XX XX  
PI Williams SR, Rothermel B;  
XX WPI; 2002-179698/23.  
DR N-PSDB; AAD30159.  
XX XX  
Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

PS Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC DNA. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 111 AA:

Query Match 51.5%; Score 537; DB 21; Length 111;  
Best Local Similarity 91.9%; Pred No. 8.7e-52;  
Matches 102; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 MPRDPSYFSSLIACVANDVSESETRAKFESLFRTYDKDTTFQYFKSKVRINFSN 60  
DB 1 MPRDPSYFSSLIACVANDVSESETRAKFESLFRTYDKDTTFQYFKSKVRINFSN 60  
OY 61 PLSADARLRLKTEFLGKEMKLYFAOQLHIGSSHLAPNPDKOFLISPPA 111  
DB 61 PLSADARLRLKTEFLGKEMKLYFAOQLHIGSSHLAPNPDKOFLISPPA 111

RESULT 13

AB71467  
ID ABB71467 standard; Protein; 292 AA.

XX ABB71467;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 41193.

XX Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2001; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PMD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL15570.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -

PS Disclosure; SEQ ID NO 41193; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutic and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (AB85737-AB87072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcc\_sequences.

XX Sequence 292 AA:

Query Match 35.7%; Score 372; DB 22; Length 292;  
Best Local Similarity 41.0%; Pred No. 9e-33;  
Matches 75; Conservative 42; Mismatches 56; Indels 10; Gaps 4;

OY 11 SSLIACVANDVSESETRAKFESLFRTYDKDTTFQYFKSKVRINFSNPLSADARL 70  
DB 116 TSLIVTHSEVPANPELKLAMEELFTFSESATFQWLRSLRVYDVAIAANARIK 175

OY 71 LKTEFLGKE-MKLYFNQTL-HIGSSHLAPNPDKOFLISPPAPVGMQVDAFPPV 128  
DB 176 LHOYEFNKKVTITCYFAQPTVPVSNKNLQPPAPVKOFLISPPAPGAPREGEPLVN 235

OY 129 YDLVATSKLQPGKYEELHAATDPTPSVVHVCESDQENEEEMERMRPRPKIITQR 188  
DB 236 HDLALALASLTGSHLHPOSBDQPAIIVHTAML-----AETGPGLOVAKP---IVQTK 287

OY 189 RPE 191  
DB 288 CPE 290

RESULT 14

AAB58452  
ID AAB58452 standard; Protein; 142 AA.

XX AAB58452;

DT 14-MAR-2001 (first entry)

XX Lung cancer associated polypeptide sequence SEQ ID 790.

XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;  
KM cardioactive; immunomodulatory; muscular active; vulnery;

XX gastrointestinal; nephrotoxic; antiinfective; gynecological;  
KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KM proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN WO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05918.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCT INC.

PA (ROSE/) ROSEN C A.

PI Ruben SM;

DR WPI; 2000-587514/55.

DR N-PSDB; AAF18328.

PT Lung cancer associated gene sequences, referred to as lung cancer  
XX antigens, useful for treatment, prevention, and diagnosis of disorders  
XX such as lung cancer -

PS Claim 11; Page 1325-1326; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and

antagonists may have neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular active general; vulnery; gastrointestinal general; nephrotic; antinefactive; gynecological; or antibacterial activity. The invention also includes antibodies specific for the protein or polynucleotide sequences. The lung cancer associated polynucleotide sequences may be used for detection of lung cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The proteins may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAF58549 are used in the course of the invention for identification and characterisation of the polynucleotide and protein sequences.

RESULT 15	
ABB29495	
ID	ABB29495 standard; Peptide; 58 AA.
XX	
AC	ABB29495;
XX	
DT	01-FEB-2002 (first entry)
XX	
DE	Peptide #2146 encoded by breast cell single exon nucleic acid probe.
XX	
XX	Human; microarray; single exon probe; gene expression; breast;
KW	disease; cancer.

nucleic acid probes for measuring gene expression in a sample derived from human breast and BR 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for determining predisposition and/or prognosing breast disease. Gene expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/obtained\\_pct\\_sequences](http://ftp.wipo.int/pub/obtained_pct_sequences).

Search completed: December 11, 2002, 11:42:39  
Job time : 31.7184 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3457 Seconds  
(without alignments)  
1117.976 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFRSFFSLIACVAND.....RPKPKIIQTRRPTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	992.5	95.2	197	1 CCPI CRIGR	O35847 cricetulus
2	912.5	87.5	198	1 CCPI MOUSE	O9jhg6 mus musculus
3	875	83.9	197	1 CCPI HUMAN	P31805 homo sapien
4	632.5	60.6	197	1 CCP2 MOUSE	O9jhg2 mus musculus
5	615.5	59.0	197	1 CCP2 HUMAN	O14206 homo sapien
6	594.5	57.0	239	1 CCP3 MOUSE	O9jkk0 mus musculus
7	592.5	56.8	241	1 CCP3 HUMAN	O9ukx8 homo sapien
8	372	35.7	292	1 NL4 DROME	O9x218 drosophila
9	351	33.7	207	1 CCPL CAEL	P53806 caenorhabdi
10	95.5	9.2	163	1 YA9F SCHPO	O09791 schizosacch
11	86.5	8.3	822	1 NAH1 CRIGR	P48761 cricetulus
12	85.5	8.2	561	1 SNTPI SCHPO	P49902 homo sapien
13	82.5	7.9	734	1 KQ08 YEAST	P36004 saccharomyc
14	82	7.9	370	1 CR42 BACSH	P05516 bacillus sp
15	81.5	7.8	1790	1 SEPA EMENI	P78621 emericella
16	81	7.8	150	1 ALI5 HEVBR	O39967 hevea bras
17	81	7.8	2329	1 CCAB HUMAN	O00975 homo sapien
18	80.5	7.7	360	1 VTPA TTVI	P19274 thermoprote
19	80	7.7	370	1 CR41 BACSH	P06575 bacillus sp
20	79.5	7.6	452	1 CN17 DICDI	P12019 dictyosteli
21	79.5	7.6	559	1 ECM1 MOUSE	O61508 mus musculus
22	79.5	7.6	587	1 RGP1 HUMAN	P46060 homo sapien
23	78.5	7.5	201	1 MAP2 SCHPO	O09180 schizosacch
24	78.5	7.5	211	1 RCN1 YEAST	P36054 saccharomyc
25	78	7.5	370	1 CR43 BACSH	P12963 bacillus sp
26	78	7.5	467	1 INVO MOUSE	P48997 mus musculus
27	78	7.5	1355	1 SALM DROME	P39770 drosophila
28	77.5	7.4	592	1 ABP1 YEAST	P15891 saccharomyc
29	77.5	7.4	627	1 DNK1 GUITH	P29215 guillardia
30	77.5	7.4	630	1 YCF2 GENVI	P31569 oenothera v
31	77.5	7.4	721	1 YCF2 GENPI	P31568 oenothera p
32	77	7.4	681	1 DP3X MYCPN	P75177 mycoplasma
33	77	7.4	691	1 CYG3 BOVIN	P19687 bos taurus

## ALIGNMENTS

RESULT 1

```

CCPI_CRIGR
ID CCPI_CRIGR STANDARD; PRT; 197 AA.
AC O35847;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Oxidative-induced protein Adapt78).
GN DSCR1 OR ADAPT78.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97329095; PubMed=9185608;
RA Crawford D.R., Leahy K.P., Abramova N., Lan L., Wang Y., Davies K.J.;
RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue
RL that is inducible by oxidative stress.";
RL Arch. Biochem. Biophys. 342:6-12(1997).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -!- INDUCTION: OXIDANT-INDUCIBLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; U60263; AB068517.1; -
SQ SEQUENCE 197 AA; 22674 MW; 0EFC68FA55826439 CRC64;

```

```

Query Match          95.2%; Score 992.5; DB 1; Length 197;
Best Local Similarity 96.5%; Pred. No. 3.5e-78;
Matches 191; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 1 MDRFRSFFSLIACVANDVDFSESETRAKFESLFRTYDKDTTFQYFKSKFRVRINFN 60
DB 1 MHFRDYNFSSLIACVANGDVFSESETRAKFESLFRTYDKDTTFQYFKSKFRVRINFN 60
OY 61 PLSAADARLRHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120
DB 61 PLSAADARLRHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120
OY 121 EDATPVINYDLLYALSKLGPCKEYELHAATDPTPSVVHVHVCESDQENEEEEEEMKRP 180
DB 121 EDATPVINYDLLYALSKLGPCKEYELHAATDPTPSVVHVHVCESDQEN-EEEEEMKRP 179
OY 181 KPKIIQTRRPTPIHLS 198

```

```

Db      180 KKKIOTRRREPTPHLS 197

|||||
|||||
|||||

RESULT 2
CCPI_MOUSE STANDARD; PRT; 198 AA.
ID_CCP1_MOUSE
AC 09JHG6: 09JKK3; 09JKK2; 09JKS1; 09UK50;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin I (Down syndrome critical region protein 1 homolog)
DB (Myocyte-enriched calcineurin interacting protein I) (MCIP1).
DS DSCR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_taxid=10090;
(1)
SEQUENCE FROM N.A. (ISOFORM A).
RP STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=20534793; PubMed=11080588;
RA Stripoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RA "The murine DSCR1-like (Down syndrome candidate region 1) gene family";
RL conserved synteny with the human orthologous genes.";
RN Gene 257:223-232(2000).
(2)
SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=20187590; PubMed=10722714;
RA Rothermel B., Vega R.B., Yang J., Wu H., Basel-Duby R.,
RA Williams R.S.;
RA "A protein encoded within the Down syndrome critical region is
RA enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725(2000).
(3)
SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.
TISSUE=Fetal brain;
RX MEDLINE=21152920; PubMed=11231093;
RA Caeas C., Martinez S., Pritchard M.A., Fuentes J.J., Nadal M.,
RA Guimera J., Arbones M., Flores J., Soriano E., Estivill X.,
RA Alcantara S.;
RA "Dscr1, a novel endogenous inhibitor of calcineurin signaling, is
RA expressed in the primitive ventricle of the heart and during
RA neurogenesis ";
RL Mech. Dev. 101:289-292(2001).
(4)
SEQUENCE FROM N.A. (ISOFORMS B AND C).
RP Fuentes J.J., Pritchard M.A., Pucharcos C., Estivill X.;
RA "Down syndrome candidate region 1 (Dscr1), one of three alternatively
RA spliced exon 1 transcripts.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
(5)
SEQUENCE FROM N.A. (ISOFORM C).
RP STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arahawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojouri T., Bono H., Kasakawa T., Sato R.,
RA Kadotani K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., Kling B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateno Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stubbli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Botfell T., Boujunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein W.J., But C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hune D.A., Kamiya N., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Monbaerts P.,
RA Norodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohneuk S.,
RA Hayashizaki Y.;
```

Query Match	Best Local Similarity	87.5%	Score 912.5	DB 1	Length 198
Matches 177, Conservative	5	Mismatches 13	Indels	3	Gaps 1
DB	1	MPRPDSYVFSSLIACVANDVFSSESRPAKEESI.PRTYDKOTTTQYPSFKRVIRINFSN	60		
DB	4	VLQIDLP--SATICHLDPRVFVGLCAKESLPRYDKOTTTQYPSFKRVIRINFSN	60		
QY	61	PLSADARLRLKTEPLGKMKLYFAQTLHGSSHLAPPNDPKOFLSPSPAPVGMKQV	120		
DB	61	PLSADARLRLKTEPLGKMKLYFAQTLHGSSHLAPPNDPKOFLSPSPAPVGMKQV	120		
QY	121	BDATPVYNDLLVAISKLGPGKXYELHAATDPTPSVYVAVCSDDQNEBEEBEMERMKRP	180		
DB	121	BDATPVYNDLLVAISKLGPGKXYELHAATDPTPSVYVAVCSDDQNEBEEBEMERMKRP	180		
QY	181	KPKIIOTRRPEYTPHLS	198		
DB	181	KPKIIOTRRPEYTPHLS	198		
RESULT 3					
CCP1_HUMAN	STANDARD	PRT	197 AA.		
AC	P53805; O00582; O00583; Q9JUN6; Q9JUF5; Q9BU69;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Calcipressin 1 (down syndrome critical region protein 1) (Myocyte-enriched calcheneurin interacting protein 1) (MCIP1) (Adapt78).				
GN	DSCR1 OR DSCR1 OR ADAPT78.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RF	SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).				
RX	MEDLINE=6121593; PubMed=8595418;				
RA	Fuentes J.-J., Pritchard M.A., Plana A.M., Bosch A., Ferrer I.,				





```

Db 71 LHETOPRCKKLTYPAOVOTPETDDKLTALAPPOPAKOFLLSPSSPPYGMWPKIDATPV 130
OY 127 INVDLLVAISKSGPGEKELHATDPTSPVVHVCESDENEEBEMRMKPKKIIQ 186
Db 131 LNYDLYVAVKLGPGEKELHAGTSTSPVVHVCDSDMEBEDRK-----TSPKPKIIQ 185
OY 187 TRRPETYP 194
Db 186 TRRPGLPP 193

RESULT 5
CCP2_HUMAN STANDARD; PRT; 197 AA.
ID CCP2_HUMAN
AC 014206:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcipressin 2 (Thyroid hormone-responsive protein ZAKI-4) (Down
DE syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE interacting protein 2) (MCIP2).
GN DSCR1L1 OR ZAKI4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RS TISSUE=fibroblast; PubMed=8662924;
RX MEDLINE=96378928;
RA Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA Yamamura H., Seo H.,
RT "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT in human skin fibroblasts";
RL J. Biol. Chem. 271:14567-14571(1996).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -1- TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC and skeletal muscle but not in placenta, lung, kidney and
CC pancreas.
CC -1- INDUCTION: By thyroid hormone.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D83407; BAA11911.1; ALT_INIT.
DR Genew; HGNC:3041; DSCR1L1.
DR MIM; 604876; -.
SQ SEQUENCE 197 AA; 21955 MW; AF39735F6661CBE0 CRC64;

Query Match 59.0%; Score 615.5; DB 1; Length 197;
Best Local Similarity 63.3%; Pred. No. 6.7e-46;
Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

OY 11 SSIACVANDVSESEETRAKFSLEFRTYDKDTTFQYFKSPKVRINFSNPLSADARLR 70
Db 11 STLVACVADVSEFTNGVCKRFGSLFRTYDDCTFQLFKFRFRRLINFSNPKSAARARIE 70
OY 71 LKATFLGKEMKLYFACTLHIGSS---HLAPENPKQPLISPPASPPVCKQVEDATPV 126
Db 71 LHETOPRCKKLTYPAOVOTPETDDKLTALAPPOPAKOFLLSPSSPPYGMWPKIDATPV 130
OY 127 INVDLLVAISKSGPGEKELHATDPTSPVVHVCESDENEEBEMRMKPKKIIQ 186
Db 131 LNYDLYVAVKLGPGEKELHAGTSTSPVVHVCDSDMEBEDRK-----TSPKPKIIQ 185

```

```

OY 187 TRRPETYP 194
Db 186 TRRPGLPP 193

RESULT 6
CCP3_MOUSE STANDARD; PRT; 239 AA.
ID CCP3_MOUSE
AC 094XKO; 09CX87;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN DSCR1L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RS STRAIN=BA1B/C; TISSUE=Brain;
RA Strippoli P., Petini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like gene family.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
(2)
RP SEQUENCE OF 40-239 FROM N.A.
RS STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischnmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Marsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okubo T., Furuno M., Aono H., Baldarelli R., Barin G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
RA Guentrich S., Hill D., Hofmann M., Hume D.A., Kamiy M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilting L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohzuki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF237889; AAF62539.1; -.
DR EMBL; AK019377; BAB31687.1; -.
DR MGD; MGI:1858220; Dscr1l2.
SQ SEQUENCE 239 AA; 27153 MW; 1B2687B47B84D272 CRC64;

Query Match 57.0%; Score 594.5; DB 1; Length 239;
Best Local Similarity 60.4%; Pred. No. 5.3e-44;
Matches 122; Conservative 19; Mismatches 48; Indels 13; Gaps 3;

OY 1 MDPRDSYVNFSSLIACVANDVSESEETRAKFSLEFRTYDKDTTFQYFKSPKVRINFSN 60

```

```

Db 40 MDLSLP---TSLFACSVHEAFVQEKERFEALFTLYDDQVTFQFKSFRRVINFSK 96
Qy 61 PLISAADARLRHLKHTFLGKEMKLYFAQTLHIG-----SSHLPAPNDKQFLISPASPVG 116
Db 97 PEAAARARIELHESFNGKLYFAQVQVGEARDKSYLLPPQTKQFLISPASPVG 156
Qy 117 WKQVEDATPVINYDLYAISKLGPGEKLYELHAATDTPSVVHVHVCESDOENEEEMER 176
Db 157 WKQSEDAMPVINYDLYCAVSKLGPGEKLYELHAGTSTPSVVHVHVCESTEEEED----- 210
Qy 177 MKRPKPKIQTTRPEYTPHLS 198
Db 211 TKNPKQKITQTRRPEAPTAALS 232

RESULT 7
CCP3 HUMAN
ID _CCP3 HUMAN STANDARD; PRT; 241 AA.
AC O9UKA8; O9UKA7; O9NUC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Calpriesin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN DSCR1L2
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Peripheral blood, and Placenta;
RX MEDLINE=2021370; PubMed=10756093;
RT "A new gene family including DSCR1 (Down syndrome candidate), region 1)
RT and ZAK1-4: characterization from yeast to human and identification of
RL DSCR1-like 2, a novel human member (DSCR1L2).";
RN Genomics 64:252-263(2000).
[2]
RP SEQUENCE OF 66-241 FROM N.A.
RA Bagguley C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (SHOWN HERE) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC kidney, liver and peripheral blood leukocytes. Lower expression in
CC all other tissues.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF176116; AAF01684.1; -
CC DR EMBL; AF176117; AAF01685.1; -
CC DR EMBL; AL034582; CAB72298.1; -
CC DR Genew; HGNC:3042; DSCR1L2.
CC DR MIM; 605860; -
CC KW Alternative splicing.
CC FT VARSPPLIC 124 133 MISSING (IN ISOFORM 2).
SQ SEQUENCE 241 AA; 27492 MW; 91AB619F5E443FBD CRC64;

Query Match 56.8%; Score 592.5; DB 1; Length 241;
Best Local Similarity 61.5%; Pred. No. 86-44;
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;

```

```

Qy 1 MDRFDSYFSSLIACVANDVDFSESETRAKFSLFRTYDKOTTFOYFKSKVRINFNS 60
Db 40 MDLSLP---TSLFACSVHEAFVQEKERFEALFTLYDDQVTFQFKSFRRVINFSK 96
Qy 61 PLISAADARLRHLKHTFLGKEMKLYFAQTLHIG-----SSHLPAPNDKQFLISPASPVG 116
Db 97 PEAAARARIELHESFNGKLYFAQVQVGEARDKSYLLPPQTKQFLISPASPVG 156
Qy 117 WKQVEDATPVINYDLYAISKLGPGEKLYELHAATDTPSVVHVHVCESDOENEEEMER 176
Db 157 WKQSEDAMPVINYDLYCAVSKLGPGEKLYELHAGTSTPSVVHVHVCESTEEEED----- 210
Qy 177 MKRPKPKIQTTRPE 191
Db 211 TKNPKQKITQTRRPD 225

RESULT 8
NLA DROME
ID _NLA DROME STANDARD; PRT; 292 AA.
AC O9XZL8; O9V391;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nebula protein.
GN NLA OR CG6072;
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA McCormack A.V.; Goldberg M.L.;
RT "Gene required for elongation of meiosis I spindle in Drosophila
RT females.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA George R.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Sutton G.G.; Wortman J.R.; Yandell M.D.; Zhang Q.; Chen L.X.;
RA Brandon R.C.; Rogers Y.-H.C.; Blazej R.G.; Champe M.; Pfeiffer B.D.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;
RA Bailly R.M.; Basu A.; Baxendale J.; Andrews-Pfannkoch C.; Baldwin D.;
RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;
RA Borkova D.; Botchan M.R.; Bouck J.; Brokstein P.; Brotter P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cawley S.; Dahlke C.; Davenport L.B.; Davies P.;
RA de Pablos B.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.; Dou P.L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferrieria S.; Fleischmann W.;
RA Fogler C.; Gabrielian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Glodek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.; Heiman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;
RA Jalali M.; Kalush F.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;
RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;
RA Lasko P.; Lei Y.; Levitsky A.A.; Li J.; Li Z.; Liang Y.; Lin X.;
RA Liu X.; Mattei B.; McIntosh T.C.; McLeod M.P.; McPherson D.;
RA Merkulov G.; Milshina N.V.; Mobarry C.; Morris J.; Moshrefi A.;
RA Mount S.M.; Moy M.; Murphy B.; Murphy L.; Muzny D.M.; Nelson D.L.;
RA Nelson D.R.; Nelson K.A.; Nixon K.; Nusskern D.R.; Pacleb J.M.;
RA Palazzolo M.; Pittman G.S.; Pan S.; Pollard J.; Puri V.; Reese M.G.;
RA Reinert K.; Remington K.; Saunders R.D.C.; Scheeler F.; Shen H.;
RA Shue B.C.; Siden-Kiamos I.; Simpson M.; Skupski M.P.; Smith T.;
RA Spier E.; Spradling A.C.; Stapleton M.; Strong R.; Sun E.;
RA Swirskas R.; Tector C.; Turner R.; Venter E.; Wang A.H.; Wang X.;

```

Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Mu D., Yang S., Yao Q.A., Ye J., Yen R.-F., Zaveloff J.S., Zhan M., Zhang G., Zhao O., Zheng L., Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,"  
 "The genome sequence of *Drosophila melanogaster*,"  
 Science 287:2185-2195(2000).  
 RL -1- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AF147700; AAD33987.1; -  
 DR EMBL: AB003712; AAF55285.1; -  
 DR FlyBase: FBgn0026629; nla.  
 SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5FE6ACF9 CRC64;  
 Query Match 35.7%; Score 372; DB 1; Length 292;  
 Best Local Similarity 41.0%; Pred. No. 7.7e-25;  
 Matches 75; Conservative 42; Mismatches 56; Indels 10; Gaps 4;  
 QY 11 SLLACVANDVSESETRAKFESLFTYDKDTTFQYKSFKRVIRNFSPNLSAADRLR 70  
 DB 116 TSLIVINSEVAPNELKAMEELFRTESESTFQWLSFRRLRVYNALIAANARK 175  
 QY 71 LKTEFLGKE-MKLYPAQTL-HIGSHLAPPNDKQFLISPPASPVGMKQVEDATPVIN 128  
 DB 176 LHOYEKNKTVICRYRQPTTPSNKNLQPPAVKQFLISPPASPVGMKQVEDATPVIN 235  
 QY 129 YDLIYAIKSGPEKYEHLAATPTPSVYVHVCESQDENEEBEEBENKRPKRIQTR 188  
 DB 236 HDLALALASLTPEGSHLHPQSEDPALIVHTML-----AETGPGIQAAP---IVQTK 287  
 QY 189 RPE 191  
 DB 288 CPE 290  
 RESULT 9  
 CCEPL\_CAEEL STANDARD; PRT; 207 AA.  
 ID CCEPL\_CAEEL STANDARD; PRT; 207 AA.  
 AC P53806; Q9U6V5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Calcipressin-1-like protein (Down Syndrome candidate region 1-like protein).  
 GN RCN-1 OR DSCR1 OR F54E7.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea; Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Br1etol N2;  
 RA Bentley D.;  
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Br1etol N2;  
 RA Strippoli P., Ienzi L., Petrini M., Carinci P., Zannotti M.;  
 RT "A new gene family including DSCR1 (Down syndrome candidate region 1) and ZAK1-4: characterization from yeast to human and identification of RT DSCR1-like 2, a novel human member (DSCR1L2).";  
 RL Genomics 64:252-263(2000).  
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

by binding to the catalytic domain of calcineurin A (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U00067; AAK20076.1; -  
 DR EMBL: AF176115; AAF01683.1; -  
 DR WormPep: F54E7.7; CE01318.  
 FT CONFLICT 205  
 FT G (IN REF. 1).  
 SQ SEQUENCE 207 AA; 23030 MW; 0154AE308AB0D5B79 CRC64;  
 Query Match 33.7%; Score 351; DB 1; Length 207;  
 Best Local Similarity 41.0%; Pred. No. 3.1e-23;  
 Matches 71; Conservative 36; Mismatches 62; Indels 4; Gaps 3;  
 QY 11 SLLACVANDVSESETRAKFESLFTYDKDTTFQYKSFKRVIRNFSPNLSAADRLR 70  
 DB 30 NALIVQPEVDFPNKDKANFSLFTQIEKDIHFDLRFRVRVIFSSPENATAKLI 89  
 QY 71 LKTEFLGKE-MKLYPAQTLH--GSHLAPPNDKQFLISPPASPVGMKQVEDATPVIN 127  
 DB 90 VQGSFPGHELKAFPAQRIWMSANSQWLSPPLEKQFLISPPASPVGMKQVEDATPVIN 149  
 QY 128 NYDLIYAIKSGPEKYEHLAATPTPSVYVHVCESQDENEEBEEBENKRP 180  
 DB 150 NFDLMARLASPAIDKYEYVHNGDELTAIVHPCEPID-VPSAIEMPRIIPR 201  
 RESULT 10  
 YV9F\_SCHPO STANDARD; PRT; 163 AA.  
 ID YV9F\_SCHPO STANDARD; PRT; 163 AA.  
 AC Q09791;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C13G6.15c in chromosome I.  
 GN SPAC13G6.15C OR SPAC24B11.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham S., Brooks K., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkart G., Aert R., Roben J., Grymoprez B., Weljens I., Vantreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Wambut R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Lelaure V., Motier S., Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
 CC by binding to the catalytic domain of calcineurin (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; Z54308; CAA91108.1; -;  
 DR EMBL; Z67757; CAA91769.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 163 AA; 18051 MW; F2575B332A276D32 CRC64;  
 Query Match 9.2%; Score 95.5; DB 1; Length 163;  
 Best Local Similarity 32.1%; Pred. No. 0.18;  
 Matches 25; Conservative 17; Mismatches 25; Indels 11; Gaps 4;  
 QY 96 LAPNPDKQFLISPPASPPVGMKQVEDATP---VINYDLYAISKLGPGKEKVELHAATDP 152  
 Db LQVPKPKWLISSPPGSPVGMKQVEDATP---VINYDLYAISKLGPGKEKVELHAATDP 152  
 QY 153 TPSVYVHVCSQDQENBE 170  
 Db LQVPKPKWLISSPPGSPVGMKQVEDATP---VINYDLYAISKLGPGKEKVELHAATDP 152  
 QY 141 GPQIVI---SEHNTKE 154  
 RESULT 11  
 NAHL\_CRIGR STANDARD; PRT; 822 AA.  
 AC P48761;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Sodium/hydrogen exchanger 1 (Na<sup>+</sup>/H<sup>+</sup>) exchanger 1 (NHE-1).  
 GN SLC9A1 OR NHE1.  
 OS Cricetulus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93192332; PubMed=8383540;  
 RA Connolly L., Pouyssegur J.;  
 RT "Nucleotide sequence of the Chinese hamster Na<sup>+</sup>/H<sup>+</sup> exchanger NHE1";  
 RL Biochim. Biophys. Acta 1172:343-345(1993).  
 CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED  
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL  
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD  
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL  
 CC TRANSDUCTION.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- PTM: PHOSPHORYLATED (POSSIBLE).  
 CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.  
 CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC  
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.  
 CC -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE  
 CC TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X68970; CAA48771.1; -;  
 DR InterPro; IPR000676; NaH\_Exchange.  
 DR Pfam; PF00999; Na\_H\_Exchange; 1.  
 DR PRINTS; PR01084; NAHEXCHNGR.  
 DR TIGRfam; TIGR00840; b\_cpai; 1.  
 KW Transmembrane; Glycoprotein; Transport; Antiport; Sodium transport;  
 KW Multigene family; Phosphorylation.  
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 13 32 A (M1) HYDROPHOBIC.  
 FT DOMAIN 33 105 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 106 127 B (M2) HYDROPHOBIC.  
 FT DOMAIN 128 130 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 131 150 C (M3) (POTENTIAL).  
 FT DOMAIN 151 162 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 163 183 D (M4) (POTENTIAL).  
 FT TRANSMEM 184 188 E (M5) (POTENTIAL).  
 FT TRANSMEM 189 210 F (M5A) (POTENTIAL).  
 FT TRANSMEM 211 230 G (M5B) (POTENTIAL).  
 FT TRANSMEM 231 251 H (M6) (POTENTIAL).  
 FT TRANSMEM 252 260 I (M7) (POTENTIAL).  
 FT TRANSMEM 261 282 J (M8) (POTENTIAL).  
 FT TRANSMEM 283 301 K (M9) (POTENTIAL).  
 FT TRANSMEM 302 322 L, HYDROPHOBIC.  
 FT TRANSMEM 323 336 M (M10) (POTENTIAL).  
 FT TRANSMEM 337 357 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT TRANSMEM 358 388 E97CIACD4EB8DAA CRC64;  
 FT TRANSMEM 389 410  
 FT TRANSMEM 411 416  
 FT TRANSMEM 417 438  
 FT TRANSMEM 439 452  
 FT TRANSMEM 453 473  
 FT TRANSMEM 474 482  
 FT TRANSMEM 483 503  
 FT TRANSMEM 504 822  
 FT CARBOHYD 374 374  
 SQ SEQUENCE 822 AA; 92003 MW; E97CIACD4EB8DAA CRC64;  
 Query Match 8.3%; Score 86.5; DB 1; Length 822;  
 Best Local Similarity 23.7%; Pred. No. 7.6;  
 Matches 40; Conservative 26; Mismatches 64; Indels 39; Gaps 8;  
 QY 44 TQYFKSKFRVIRNFSPNLSAA--DARLRLHKTFLGKMKLYFAQTLL----- 90  
 Db 646 TRQLRSYNRHTL-VADPYEAWNQMLLRQKARQLEQKMSNYLTPAKKLDSPMSRAR 704  
 QY 91 IGSSHLA-PPNDKQFLISPPASPPVGMKQVEDATPVIYDLYAISKLGPGKEKVELHAA 149  
 Db 705 IGSDPLAVEPKADLPVITIDPASP-----QSPESVDLVNEELKAKV-----LGVN 749  
 QY 150 TDPTFSVVHVHVCSDQENBEEMERMKRKPCKI-----IQTRRPEYTP 194  
 Db 750 RDPT-----RLTRGEDEDEDEGVIIMRRKEPSSPGTDVFTPAWYSP 793  
 RESULT 12  
 SNTC\_HUMAN STANDARD; PRT; 561 AA.  
 ID SNTC\_HUMAN  
 AC P49902;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cytosolic purine 5'-nucleotidase (EC 3.1.3.5).  
 GN NT5B OR NT5CP OR PNT5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

```

RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95091836; PubMed=7999131;
RA Oka J., Matsumoto A., Hosokawa Y., Inoue S.;
RT "Molecular cloning of human cytosolic purine 5'-nucleotidase.";
RN Biochem. Biophys. Res. Commun. 205:917-922(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Straube R.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY HAVE A CRITICAL ROLE IN THE MAINTENANCE OF A
CC CONSTANT COMPOSITION OF INTRACELLULAR PURINE/PYRIMIDINE
CC NUCLEOTIDES IN COOPERATION WITH OTHER NUCLEOTIDASES.
CC PREFERENTIALLY HYDROLYSES INOSINE 5'-PRIME-MONOPHOSPHATE (IMP) AND
CC OTHER PURINE NUCLEOTIDES.
CC -1- CATALYTIC ACTIVITY: A 5'-ribonucleotide + H(2)O = a ribonucleoside
CC + phosphate.
CC -1- ENZYME REGULATION: ALLOSTERICALLY ACTIVATED BY VARIOUS COMPOUNDS,
CC INCLUDING ATP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DISEASE: ABERRATION OF THIS ENZYME MAY BE RELATED TO THE URATE
CC PRODUCTION IN HYPERURICEMIA AND GOUT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38524; BA007529.1; -
DR EMBL; BC001595; AA001595.1; -
DR Genew; HGNC:8022; NT5C2.
DR MIM; 600417; -
KW Hydroxylase; Allosteric enzyme.
FT DOMAIN 549 561 ASP/GLU-RICH (ACIDIC).
FT SEQUENCE 561 AA; 64969 MW; 4C27D762575DE0EA2 CRC64;
SQ
Query Match 8.2%; Score 85.5; DB 1; Length 561;
Best Local Similarity 24.2%; Pred. No. 5.8;
Matches 47; Conservative 29; Mismatches 71; Indels 47; Gaps 10;
OY 11 SLLACVANDVDFSESETRAKFSLPR-----TYDKOTPE-----QYFKSPK 52
DB 386 SLLFELQSLDIFLAEVLYHLDSSNERPDISIORIKKVTHTDMDCYGMNGSLFRSGS 445
OY 53 RVRINFSNLSAADARLRLHKTETFLGKEMKLYFAOTLHIGSHLAPPND-----DKQ 104
DB 446 RQTLRPSQVMRYAD-----LYASFIN-----LVYRPFYLPRAHVLVMPHSTVEHTVDIN 498
OY 105 FLISPPAS---PVGKQVEDATPVINVDLLVIAISKLGSEKELMAADPPSVVHVC 161
DB 499 EMEPSPLATNRTSVDFKD-----TDYKRHOLTRISISIKPNNLPL-----APOELTH-C 547
OY 162 ESDQNEEEEEEME 175
DB 548 HDEDDDEEEEEEE 561

```

```

CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=94378719; PubMed=8091858;
RA Vandendol W., Bolle P.-A., Dion C., Portetelle D., Hilger F.;
RT "Sequencing and analysis of a 20.5 kb DNA segment located on the left
RT arm of yeast chromosome XI.";
RN Yeast 10:525-S33(1994).
RN -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z26878; CA81519.1; -
DR EMBL; Z28168; CA82010.1; -
DR PIR; S37998; S37998.
DR PIR; S38413; S38413.
DR PIR; S44593; S44593.
DR SGD; S0001651; YKL168C.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 422 722
FT NP_BIND 428 436 ATP (BY SIMILARITY).
FT BINDING 465 465 ATP (BY SIMILARITY).
FT ACT_SITE 573 573 BY SIMILARITY.
FT SEQUENCE 734 AA; 83656 MW; 896D26B060A917D CRC64;
SQ
Query Match 7.9%; Score 82.5; DB 1; Length 734;
Best Local Similarity 21.4%; Pred. No. 15;
Matches 42; Conservative 29; Mismatches 64; Indels 61; Gaps 8;
OY 11 SLLACVANDVDFSESETRAKFSLPR-----TYDKOTPE-----FOYFKSPK 53
DB 155 SSVSCDSNSNGTSSSDQWAMDLSDDNDLTPYRGSNKDILSKDRAPVNYIDYXK 214
OY 54 VRI---NFSNPLSAADARLRLHKTETFLGKEMKLYFAOTLHIGSHLAPPNDKQFLIS 109
DB 215 XALRATSYVPLPS-----KQYNERLYTRSH-----PDESLSL 252
OY 110 P---ASPPY-----GKQVEDATPVINVDLLVIAISKLGSEKELH---AATDTPSVV 157
DB 253 PFPAQADVQCIIEONGFVYEDGSHENIKLSGVIAKLEKNSLPRAHQSGLSRRLGIT 312
OY 158 V-----HVESDDQEN 167
DB 313 LSGLFKHKHNECDTEN 328

```

```

RESULT 13
K08_YEAST STANDARD; PRT; 734 AA.
ID_K08_YEAST
AC P36004;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase YKL168C (EC 2.7.1.-).
GN YKL168C OR YKL632.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

```

RESULT 14
CR42_BACSH STANDARD; PRT; 370 AA.
ID_CR42_BACSH
AC P05516;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 41.9 kDa insecticidal toxin (Strain 2297).
GN B1N4.
OS Bacillus sphaericus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

```

```

OX NCBI_TaxID=1421;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88234018; PubMed=3375083;
RA Handley J., Berry C.;
RT "Bacillus sphaericus strain 2297: nucleotide sequence of 41.9 kDa
  toxin gene";
RL Nucleic Acids Res. 16:4168-4168(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88197989; PubMed=3360740;
RA Baumann L., Broadwell A.H., Baumann P.;
RT "Sequence analysis of the mosquitoicidal toxin genes encoding 51.4-
  and 41.9-kilodalton proteins from Bacillus sphaericus 2362 and
  2297.";
RL J. Bacteriol. 170:2045-2050(1988).
CC CC -!- FUNCTION: THIS TOXIN IS ACTIVE AGAINST CULEX AND ANOPHELES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07025; CAA30074.1;
DR PIR; S01856; S01856.
DR PIR; D28211; D28211.
KW Toxin; Sporulation.
SQ SEQUENCE 370 AA; 41917 MW; 545D669505911D9 CRC64;

Query Match 7.9%; Score 82; DB 1; Length 370;
Best Local Similarity 21.1%; Pred. No. 6.9;
Matches 40; Conservative 27; Mismatches 79; Indels 44; Gaps 7;

QY 1 MDRFDSYFNSSLIACVANDVFESE--SETRAKFESLFRFYD----- 40
DB 21 MDPVNSYFPCIHAPSAPNGDITEICSRNNQYFFPDDGGRVIANRHNSVFTGEA 80
QY 41 -----KDTTFQYKSKFRVRINF-----SNPLSADARLRHLKTEFLGKEMKLYPAQT 88
DB 81 TSVVSDIYTGSPLOQFFREFKRTNSTYLAIQNPESATDVPALEPNSHEL--PSRLFTNN 138
QY 89 LHIGSSHLAPNDKQFLISPPASPVGWQVEDATPVINDLLYAISKLGPGKRYELHA 148
DB 139 IENNSNLI--SNKEQIYLTPLSLPE---NEQYKPTPLVS-----GDDIGPNQSEKSI 188
QY 149 ATDPTPSVVV 158
DB 189 GSTLIPLCIMV 198

RESULT 15
SEPA EMENI
ID SEPA EMENI STANDARD; PRT; 1790 AA.
AC P78621; Q00760;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytokinesis protein sepa (Phi/2 protein) (Forced expression
  inhibition of growth A).
GN SEPA OR FIGA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97361829; PubMed=9218790;
RA Harris S.D., Hamer L., Sharpless K.E., Hamer J.E.;
RT "The Aspergillus nidulans sepa gene encodes an Fhl/2 protein involved
  in cytokinesis and the maintenance of cellular polarity.";

```

```

RL EMBO J. 16:3474-3483(1997).
RN [2]
RP REVISIONS TO 143-153; 207; 1071-1109 AND 1644.
RA Hamer L., Harris S.D., Sharpless K.E., Hamer J.E.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 559-1790 FROM N.A., AND FUNCTION.
RC STRAIN=FGSC 26;
RX MEDLINE=95229045; PubMed=7713416;
RA Marthou J.F., Adams T.H.;
RT "Identification of developmental regulatory genes in Aspergillus
  nidulans by overexpression.";
RL Genetics 139:537-547(1995).
CC CC -!- FUNCTION: INVOLVED IN CYTOKINESIS. OVEREXPRESSION RESULTS IN
  GROWTH INHIBITION.
CC CC -!- DOMAIN: DRFS ARE REGULATED BY INTRAMOLECULAR GBD-DAD BINDING WHERE
  RHO-GTP ACTIVATES THE DRFS BY DISRUPTING THE GBD-DAD INTERACTION
  (BY SIMILARITY).
CC CC -!- SIMILARITY: CONTAINS 1 GTPASE-BINDING DOMAIN (GBD).
CC CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 1 (FHL) DOMAIN.
CC CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.
CC CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 3 (FH3) DOMAIN.
CC CC -!- SIMILARITY: CONTAINS 1 DRF AUTOREGULATORY DOMAIN (DAD).
CC CC -!- SIMILARITY: BELONGS TO THE FORMIN HOMOLOG FAMILY. BNI1
  SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83658; AAB63335.3;
DR EMBL; L36341; AAA33306.1; ALT_SEQ.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
KW Cell division; Coiled coil.
FT DOMAIN 258 486 GBD.
FT DOMAIN 376 718 FH3.
FT DOMAIN 724 811 COILED COIL (POTENTIAL).
FT DOMAIN 955 1136 FH1 (PRO-RICH).
FT DOMAIN 1141 1658 FH2.
FT DOMAIN 1435 1568 COILED COIL (POTENTIAL).
FT DOMAIN 1586 1600 DAD.
FT DOMAIN 1608 1611 ARG/LYS-RICH (BASIC).
FT CONFLICT 1071 1109 PPPPPPGGFGGPPPPPPPGGFGGPPPPPPPPPPPGGARG
  -> LPSPSSTGWFMTTAAASPTWFWRSTSAASAYWCI
  C (IN REF. 3).
FT CONFLICT 1476 1476 D -> V (IN REF. 3).
FT CONFLICT 1504 1504 V -> L (IN REF. 3).
FT CONFLICT 1504 1504 V -> L (IN REF. 3).
SQ SEQUENCE 1790 AA; 197355 MW; 192136DE2EF2A75B CRC64;

Query Match 7.8%; Score 81.5; DB 1; Length 1790;
Best Local Similarity 29.9%; Pred. No. 54;
Matches 26; Conservative 11; Mismatches 45; Indels 5; Gaps 2;

QY 98 PNPDKQLISPPASPP-----VGWQVEDATPVINDLLYAISKLGPGKRYE-LHAATDP 152
DB 1101 PPPPGGARGVPPPPPPPGTGVIGWRANYLASQAPSHAIPVWSSIRPKKALKALHWDKVD 1160
QY 153 TPVVVHVVCESDQNEEEEEEMERMKR 179
DB 1161 TPQVTVWATHGTTPOEKEKYVELAKR 1187

Search completed: December 11, 2002, 11:39:41
Job time : 9.3457 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 97.8236 Seconds  
(without alignments)  
13789.580 Million cell updates/sec

Title: US-09-782-953-1  
Perfect score: 599  
Sequence: 1 9aggtgcaaggaacctcca.....gaggaggaaggaagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 101002;\*

1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*

2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*

3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*

4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*

5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*

6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*

7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*

8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*

9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*

10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*

11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*

12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*

13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*

14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*

15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*

16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*

17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*

18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*

19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*

20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*

21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*

22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*

23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*

24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599	100.0	599	24	Human MCIP associa
2	504.4	84.2	597	24	Mouse MCIP associa
3	426.2	71.2	597	24	Human MCIP associa
4	413.8	69.1	2331	24	Human MCIP associa
5	404.8	67.6	2348	24	Human DNA sequence
6	357.8	59.7	2358	21	Lung cancer associ
7	349.2	58.3	2212	24	Human MCIP associa
8	341.8	57.1	2173	24	Gene #3341 used to
9	340.2	56.8	2174	20	Human DSCR1 coding

10	264.4	44.1	442	21	AAC01774	Human secreted pro
11	202.2	33.8	3184	24	AAD30157	Human MCIP associa
12	202.2	33.8	3184	24	ABL651768	Colon adenocarcino
13	202.2	33.8	3184	24	ABL651768	Lung cancer relate
14	197.6	33.0	934	22	AAF25338	Nucleotide sequenc
15	197.6	33.0	3159	24	ABA91457	Human Down syndrom
16	191	31.9	594	24	AAD30154	Mouse MCIP associa
17	189.8	31.7	720	24	AAD30159	Human MCIP associa
18	188.4	31.5	828	24	AAD30158	Human MCIP associa
19	181.8	30.4	615	24	ABA91463	Rat Down syndrome
20	153	25.5	1021	24	ABA91463	Rat Down syndrome
21	151.6	25.3	640	24	ABA91458	Human colon cancer
22	144.6	24.1	412	22	ABA56172	Human foetal liver
23	144.6	24.1	412	22	AAK04364	Human brain expres
24	144.6	24.1	412	22	AAK04364	Human brain expres
25	144.6	24.1	412	22	AAI35821	Probe #4382 for ge
26	144.6	24.1	412	22	AAI35821	Probe #4507 used t
27	144.6	24.1	412	24	ABSO4419	Probe #4263 used t
28	144.6	24.1	446	22	ABA43406	Human genome-deriv
29	144.6	24.1	446	22	ABA43406	Human breast cell
30	144.6	24.1	446	22	ABA53852	Human foetal liver
31	144.6	24.1	446	22	ABA53852	Probe #2064 for ge
32	144.6	24.1	446	22	AAK02113	Human brain expres
33	144.6	24.1	446	22	AAK27563	Human bone marrow
34	144.6	24.1	446	22	AAI12148	Probe #2081 for ge
35	144.6	24.1	446	22	AAI33493	Probe #2179 used t
36	144.6	24.1	446	22	AAI02061	Probe #2052 used t
37	91.8	15.3	111	22	ABSO2041	Human genome-deriv
38	91.8	15.3	111	22	ABA68757	Human foetal liver
39	91.8	15.3	111	22	AAK17102	Human brain expres
40	91.8	15.3	111	22	AAI23650	Probe #13583 for g
41	91.8	15.3	111	22	AAI48960	Probe #17646 used
42	91.8	15.3	111	22	AAI09265	Probe #9256 used t
43	77.4	12.9	123	22	ABSO1690	Human genome-deriv
44	77.4	12.9	123	22	ABA47285	Human breast cell
45	77.4	12.9	123	22	ABA65170	Human foetal liver
						Probe #10738 for g

## ALIGNMENTS

RESULT 1  
AAD30151  
ID AAD30151 standard; DNA; 599 BP.

AC AAD30151;

XX 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #1.

DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; human; ds.

XX Homo sapiens.

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX (WILL//) WILLIAMS S R.

XX (ROTH//) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

PS Disclosure; Page 147; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.

XX Sequence 599 BP; 164 A; 150 C; 149 G; 136 T; 0 other;

Query Match 100.0%; Score 599; DB 24; Length 599;  
 Best Local Similarity 100.0%; Pred. No. 7,46-104;  
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGTGCAAGGAACTCCAGCTGGGCTTGAGTGAAGAGAGAGAGCTGCTTAAGCT 60  
 DB 1 GAGGTGCAAGGAACTCCAGCTGGGCTTGAGTGAAGAGAGAGAGCTGCTTAAGCT 60  
 OY 61 CTGCCCCGTGAAAGACAGATGATTTTAAAGGACTTTAGCTCAATTTAGTCCCTGAT 120  
 DB 61 CTGCCCCGTGAAAGACAGATGATTTTAAAGGACTTTAGCTCAATTTAGTCCCTGAT 120  
 OY 121 TCGTGTGTGGCAAGAGATGATGTTCTTACGCGAAAGTGAGACAGAGGCCAAATTTGAATC 180  
 DB 121 TCGTGTGTGGCAAGAGATGATGTTCTTACGCGAAAGTGAGACAGAGGCCAAATTTGAATC 180  
 OY 181 CCTCTTCAGAACATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 DB 181 CCTCTTCAGAACATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 OY 241 CCGGATMAAATTGACAGACCCCTTATCTGAGCCGATGCCAGGCTCGGCTGCACAAAGC 300  
 DB 241 CCGGATMAAATTGACAGACCCCTTATCTGAGCCGATGCCAGGCTCGGCTGCACAAAGC 300  
 OY 301 CCGATTCCTGGGGAAGGAATGAAGTTGATTTTGTCTGAGACTTTACACATAGGAAGTTC 360  
 DB 301 CCGATTCCTGGGGAAGGAATGAAGTTGATTTTGTCTGAGACTTTACACATAGGAAGTTC 360  
 OY 361 ACACCTGGGCTCCGCAATCCCGCAAAAGATTCTCATCTCCCTCCGCTCTCTCC 420  
 DB 361 ACACCTGGGCTCCGCAATCCCGCAAAAGATTCTCATCTCCCTCCGCTCTCTCC 420  
 OY 421 GTTGCTGGAACAGTAGAAGATGCCACCCGCTGATTAATTCATTCATTTATATGCC 480  
 DB 421 GTTGCTGGAACAGTAGAAGATGCCACCCGCTGATTAATTCATTCATTTATATGCC 480  
 OY 481 ATTCTCAAGCTGGGGCCAGAGAGAGATGAATCTGATGACAGAGAGAGAGAGAGAGAG 540  
 DB 481 ATTCTCAAGCTGGGGCCAGAGAGAGATGAATCTGATGACAGAGAGAGAGAGAGAGAG 540  
 OY 541 AGTGTGTGTGTCACAGTGTGTGAGAGTGAACCAAGAGATAGAGAGAGAGAGAGAGAT 599  
 DB 541 AGTGTGTGTGTCACAGTGTGTGAGAGTGAACCAAGAGATAGAGAGAGAGAGAGAT 599

RESULT 2  
 AAD30153  
 ID AAD30153 standard; DNA; 597 BP.  
 XX  
 AC AAD30153;  
 XX

DT 17-MAY-2002 (first entry)  
 XX Mouse MCIP associated DNA #3.  
 DE

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; mouse; gene; ds.  
 XX

XX Mus musculus.

XX Key Location/Qualifiers  
 FH 1.597  
 FT CDS /tag= a  
 FT /product= "Mouse MCIP associated protein #3"

XX W0200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001MO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.

PI Williams SR, Rothermel B;

XX WPI: 2002-178698/23.  
 DR P-PSDB: AAE18912.

PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -

PS Disclosure; Page 150-151; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated DNA  
 CC Note: This sequence has been described as human MCIP3 encoding DNA  
 CC in the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.

XX Sequence 597 BP; 170 A; 156 C; 142 G; 129 T; 0 other;

Query Match 84.2%; Score 504.4; DB 24; Length 597;  
 Best Local Similarity 99.6%; Pred. No. 3,76-153;  
 Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 83 GATTTTAAAGGAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 142  
 DB 4 GATTTTAAAGGAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 63

OY 143 GTCTTCAAGGAAAGTGAAG 202  
 DB 64 GTCTTCAAGGAAAGTGAAG 123

OY 203 GACACCACTTCAGATTTTAAAGAGCTTCAAAAGTGTCCGATTAACCTTCAGAAACCC 262  
 DB 124 GACACCACTTCAGATTTTAAAGAGCTTCAAAAGTGTCCGATTAACCTTCAGAAACCC 183



PN WO000204491-A2.  
 XX 17-JAN-2002.  
 PD 06-JUL-2001; 2001WO-US21662.  
 PE 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA.) UNIV TEXAS SYSTEM.  
 PA (MILL.) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 PI Williams SR, Roethermel B;  
 XX WPI; 2002-179698/23.  
 DR P-PSDB; MAE18914.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX Claim 72; Page 155-157; 174pp; English.  
 PS The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding; expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC Note: This sequence has been described as a promoter in claim 72 of  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.  
 XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;  
 SQ  
 Query Match 69.1%; Score 413.8; DB 24; Length 2331;  
 Best Local Similarity 85.1%; Pred. No. 1.9e-123;  
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
 QY 39 GAGCGAGTCGTTGCTTAAGCGTGTGCCCCGTAAGAAAGCAGATGATTTAGGAGCTTTA 98  
 DB 103 GACTGGCGTGGTGTGAGCCGCTTCACTGTAAGAAAGCAGATGATTTAGAACTTTA 162  
 QY 99 GCTACATTTTAACTCCCTGATGCTGTGTGCAAAAGATGATGCTTTCAGCGAAAGTG 158  
 DB 163 ACTACAGTTTAACTCCCTGATGCTGTGTGCAAAAGATGATGCTTTCAGCGAAAGTG 222  
 QY 159 AGACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGACACCACTTCCAGT 218  
 DB 223 AAACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGACACCACTTCCAGT 282  
 QY 219 ATTTTAAGAGCTTCAAAAGCTGTCCGATTAATCTTGCAACCCCTTATCGACGGGATG 278  
 DB 283 ATTTTAAGAGCTTCAAAAGCTGTCCGATTAATCTTGCAACCCCTTATCGACGGGATG 342  
 QY 279 CGAGGCTGGCGGTGACAAAGACCGAGTTCTGGGGAAGAAATGAATGTTATTTGCTC 338  
 DB 343 CGAGGCTGGCGGTGACAAAGACCGAGTTCTGGGGAAGAAATGAATGTTATTTGCTC 402  
 QY 339 AGACTTTACACATAGGAAGTTTACACCTGCTCCGCC-AAATCCGACAAACAGTTTCTCA 397  
 DB 403 AGACTTTACACATAGGAAGTTTACACCTGCTCCGCC-AAATCCGACAAACAGTTTCTCA 462  
 QY 398 TCTCCCTCCGGGCTCTCTCTCCCTTGGGTAAGAAAGATGAAAGTCCACCCCGCTCA 457

DB 463 TCTCCCTCCGGGCTCTCTCTCCCTTGGGTAAGAAAGATGAAAGTCCACCCAGTCA 522  
 QY 458 TAAATTACGATCTTTATATGATCCATTCGAAAGCTGGGCGCAGAGAGAGATATGACTGC 517  
 DB 523 TAAACTATGATCTCTTATATGATCCATTCGAAAGCTGGGCGCAGAGAGAGATATGACTGC 582  
 QY 518 ATGACGACGACAGACACCACTCCGAGTGTGTGTCACGTCGTGAGAGTACCAAGAGA 577  
 DB 583 AGCGAGCGACTGACACCACTCCGAGTGTGTGTCACGTCGTGAGAGTATGATGATGAGAGA 642  
 QY 578 ATGAGAGAGAGAGAGAA 594  
 DB 643 AGAGAGAGAGAGAGAA 659  
 RESULT 5  
 ID AAS94805 standard; DNA; 2348 BP.  
 XX AAS94805;  
 AC AAS94805;  
 DT 14-FEB-2002 (first entry)  
 XX Human DNA sequence #60 expressed during foam cell differentiation.  
 DE Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
 XX cardiovascular disorder; coronary artery disease; gene therapy; db.  
 XX Homo sapiens.  
 OS WO200177389-A2.  
 PD 18-OCT-2001.  
 XX 04-APR-2001; 2001WO-US11128.  
 PF 05-APR-2000; 2000US-195106P.  
 PR (INCY-) INCYTE GENOMICS INC.  
 PA Shiffman D, Somogyi R, Lawn R, Sellhammer JF, Porter GJ, Mikita T;  
 PI Tai J;  
 XX WPI; 2002-010925/01.  
 DR Composition useful for diagnosis of conditions, disorders or diseases  
 XX associated with atherosclerosis, comprises several polynucleotides that  
 XX are differentially expressed in foam cell development -  
 PS Claim 1; Page 112-113; 315pp; English.  
 XX The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a  
 CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the  
 CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used  
 CC as PCR primers and probes. The polynucleotide sequences of the invention are also  
 CC useful in gene therapy. AAS9476-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation.  
 XX Sequence 2348 BP; 638 A; 473 C; 550 G; 687 T; 0 other;  
 SQ  
 Query Match 67.6%; Score 404.8; DB 24; Length 2348;  
 Best Local Similarity 84.1%; Pred. No. 1.6e-120;  
 Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;  
 QY 39 GAGCGAGTCGTTGCTTAAGCGTGTGCCCCGTAAGAAAGCAGATGATTTAGGAGCTTTA 98

Db 110 GACTGGTGGGCTCTAGCGCTTTCACCTGTAGAAAGCAAGATGCAATTTAGAAATTTA 169  
Qy 99 GCTACATTTTAGCTCCCTGATTTGCTGTGGCAACGATGATGCTTTCAGCGAAAGTG 158  
Db 170 ACTACAGTTTTAGCTCCCTGATTTGCTGTGGCAACGATGATGCTTTCAGCGAAAGTG 229  
Qy 159 AGACCAGGGCCAAATTTCAATCCCTTTCAGAACATATGACAAAGACACACCTTCAGT 218  
Db 230 AAACCCAGGGCCAAATTTGAGTCCCTTTCAGAACATATGACAAAGACATCACCTTTCAGT 289  
Qy 219 ATTTTAAGAGCTTCAAAAGCTGTCCGGATAAACTTCAGCAACCTTCATCTCGACCGCATG 278  
Db 290 ATTTTAAGAGCTTCAAAAGCTGTCCGGATAAACTTCAGCAACCTTCATCTCGACCGCATG 349  
Qy 279 CCAGCTCGGCTGCAACAGACCGAGTTCCTGGGAAAGGAATCAAGTTGATTTTTCCTC 338  
Db 350 CCAGCTCGGCTGCAACAGACCGAGTTCCTGGGAAAGGAATCAAGTTGATTTTTCCTC 409  
Qy 339 AGACTTTACATAGGAGTTTCACACCTGG--CTCCGCCAATCCCGAACAACAGTTCCTC 396  
Db 410 AGACTTTACATAGGAGTTTCACACCTGGCTCCGCCAATCCCGAACAACAGTTCCTC 469  
Qy 397 ATCTCCCTCCGGCTCTCTCCCTGGTGGTGGAAACAAGTAGAAGATGCCACCCCGTC 456  
Db 470 ATCTCCCTCCGGCTCTCCGGCAGTGGATGGAACAAGTGGACCCCGCAGTC 529  
Qy 457 ATAAATTACGATCTTTATATGCTATCCAGTCCAGTGGGCGCAGAGAGAGTATGAATG 516  
Db 530 ATAAATTACGATCTTTATATGCTATCCAGTCCAGTGGGCGCAGAGAGAGTATGAATG 589  
Qy 517 CATGACGACAGACACCACTCCACGATGGTGGTGGAAACAAGTAGAAGATGCCACCCCGTC 576  
Db 590 CACGACGAGTACACCACTCCACGATGGTGGTGGAAACAAGTGGACCCCGCAGTC 649  
Qy 577 AATGAGGAGGAAGAGAA 594  
Db 650 AAGGAGGAGGAAGAGAA 667

RESULT 6  
AAF18328  
ID AAF18328 standard; DNA; 2358 BP.  
AC AAF18328;  
DT 14-MAR-2001 (first entry)  
XX Lung cancer associated polynucleotide sequence SEQ ID 347.  
DE Human; lung cancer associated protein; neuroprotective; cytostatic;  
KW cardioactive; immunomodulatory; muscular active; vulnerary;  
KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
XX proliferative disorder; wound healing; infectious disease; ds.  
OS Homo sapiens.  
XX WO200055180-A2.  
PN 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05918.  
PF 12-MAR-1999; 99US-0124270.  
PR (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX Ruben SM;  
PI WPI; 2000-587514/55.  
DR P-PSDB; AAB58452.  
XX

PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer -  
XX  
PS Claim 1; Page 805-806; 1425pp; English.  
XX  
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective, cytostatic; cardioactive;  
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.  
XX  
SQ Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

Query Match 59.7%; Score 357.8; DB 21; Length 2358;  
Best Local Similarity 80.5%; Pred. No. 2.9e-105;  
Matches 416; Conservative 2; Mismatches 99; Indels 0; Gaps 0;  
Qy 78 AGAATGATTTTAGGACCTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGGCAAAAGC 137  
Db 112 AGATGGAGGAGTGAGCTGCGAGACCTGCCAGCGCCACCATCCCTGTCACCTGGACC 171  
Qy 138 ATGATGCTTTCAGGAAAGTGAGACACGAGGCGCAAAATTTGAATCCCTCTTCAAGACATATG 197  
Db 172 CGCGCTGTTCGTGACGGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGACGTATG 231  
Qy 198 ACAAGGACACCACTTCAGTATTTTAAGAGCTTCAAAAGTGTCCGGATAAACTTCAGCA 257  
Db 232 ACAAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAAGGAGTCAAGATAAACTTCAGCA 291  
Qy 258 ACCCTTTATCTGACGCGATGCCAGGCTGCGGCTGCACAGACCGAGTTCCTGGGAAAGG 317  
Db 292 ACCCTTTCTCGGACGAGATGCCAGGCTCCAGTGTCAATAGACTGATTTCTCGGAAAGG 351  
Qy 318 AATGAAGTTGATTTTGTCTCAGACTTTACATAGGAAGTTCAACCTCGCTCCGCAA 377  
Db 352 AATGAAGTTTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTCGCTCCGCAA 411  
Qy 378 TCCGCAACAAAGTTCCTCATCTCCCTCCGCTCTCTCCGCTGTGGTGGGAAACAAGT 437  
Db 412 TCCAGCAAGAGTTTCTGATCTCCCTCCGCTCTCTCCGCTGTGGTGGGAAACAAGT 471  
Qy 438 AGAAGATGCCACCCCTCATAAATACGATCTTTTATATGCCATCTCCAAGCTGGGCC 497  
Db 472 GGAAGATGCCACCCCTCATAAATACGATCTTTTATATGCCATCTCCAAGCTGGGCC 531  
Qy 498 AGGAGAAATGAACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 557  
Db 532 AGGGGAAAGTATGAATTTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 591  
Qy 558 GTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGAGAA 594  
Db 592 ATGTAGAGTGATCAAGAGGAAGGAGGAAGAGAGAA 628

RESULT 7  
AAD30156  
ID AAD30156 standard; DNA; 2212 BP.  
XX  
AC AAD30156;

XX 17-MAY-2002 (first entry)  
 DT Human MCIP associated DNA #2.  
 DE  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 PH 25..618  
 FT CDS /\*tag= a  
 FT /product= "Human MCIP associated protein #2"  
 FT  
 PN WO200204491-A2.  
 PD 17-JAN-2002.  
 PF 06-JUL-2001; 2001WO-US21662.  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIT TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 PI Williams SR, Rothermel B;  
 XX WPI: 2002-179698/23.  
 DR P-PSDB; AAB18915.  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 PT  
 PS Example 1; Page 159-161; 174pp; English.  
 XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC Note: This sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.  
 CC  
 XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;  
 SQ  
 XX  
 Query Match 58.3%; Score 349.2; DB 24; Length 2212;  
 Best Local Similarity 80.9%; Pred. No. 1.8e-102;  
 Matches 419; Conservative 0; Mismatches 98; Indels 1; Gaps 1;  
 OY 78 AGAATGATTTAGGACCTTACATTTAGCTCCCTGATTCCTGTGCGAAGC 137  
 DB 23 AGATGAGAGAGGTGACCTCGACGACCTGCCAGGCCACATGCGCTGACCTGGACC 82  
 OY 138 ATGATGCTTCAGCGAAGTAGACAGCGCCAAATTTGAATCCTCTTCGAAACATATG 197  
 DB 83 CGCGGTGTGTGTGACGCGCTGTGCCGCGCAATTTGAGTCCCTCTTTAGGAGATG 142  
 OY 198 ACAAGACACACCTTCCTGATTTTAAAGAGCTTCAACAGTGTCCGATTAACCTCAGCA 257  
 DB 143 ACAAGACATCACCTTCCTGATTTTAAAGAGCTTCAACAGATCAGAAATTAACCTCAGCA 202

OY 258 ACCCTTATCTGAGCCGATGCGAGCTGCGCTGCAACAAGCCGATTCCTGGGAAAG 317  
 DB 203 ACCCTTCTCCGACGAGATGCGAGCTCCAGCTGCATAGACTGATTCCTGGGAAAG 262  
 OY 318 AATGAAAGTTGATTTTTCCTCAGACTTTACATAGAAAGTTCCAGCTGCGCC-A 376  
 DB 263 AATGAAAGTTATTTTTCCTCAGACTTTACATAGAAAGTTCCAGCTGCGCCAA 322  
 OY 377 ATCCGACAAACAGTTCCCTCATCTCCCTCCGACCTCTCCCTCCGCTGGAAACAG 436  
 DB 323 ATCCGACAAACAGTTCTCATCTCCCTCCGACCTCTCCGACGAGTGGATGAAACAG 382  
 OY 437 TAGAAGATGCCACCCCGCTATTAATTAATTCATCTTTATTTGCAATCTCCAACTGGGGC 496  
 DB 383 TGGAAGATGCCACCCCGCTATTAATTAATTCATCTTTATTTGCAATCTCCAACTGGGGC 442  
 OY 497 CAGAGAGAAAGTATGAAGTGCATGACGACGACCACTCCAGTGTGTGATGACG 556  
 DB 443 CAGAGAGAAAGTATGAAGTGCATGACGACGACCTGACACCACTCCAGGTGTGTGATG 502  
 OY 557 TGTGTGAGAGTGAACCAAGATGAGAGAGAGAGCA 594  
 DB 503 TATGTGAGAGTGAACCAAGATGAGAGAGAGAGCA 540  
 RESULT 8  
 ABN96843  
 ID ABN96843 standard; DNA; 2173 BP.  
 XX  
 AC ABN96843;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Gene #3341 used to diagnose liver cancer.  
 XX  
 KW Gene: liver cancer; ds; hepatocellular carcinoma; hepatocytic;  
 KW metastatic liver tumour; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200229103-A2.  
 PN  
 XX 11-APR-2002.  
 PD  
 XX 02-OCT-2001; 2001WO-US30589.  
 PF  
 XX 02-OCT-2000; 2000US-237054P.  
 PR  
 XX (GENE-) GENE LOGIC INC.  
 PA  
 XX Horne D, Alvares C, Pires-da-Silva S, Vockley JG;  
 PI WPI: 2002-426119/45.  
 DR  
 XX  
 PT Diagnosing and detecting the progression of liver cancer.  
 PT hepatocellular carcinoma or metastatic liver tumor in a patient.  
 PT involves detecting the level of expression of two or more genes in a  
 PT liver tissue sample -  
 PT  
 PS Claim 1; SEQ ID NO 3341; 298bp; English.  
 XX  
 CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumour in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatocytic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;

Query Match 57.1%; Score 341.8; DB 24; Length 2173;  
Best Local Similarity 88.8%; Pred. No. 4.4e-100;  
Matches 381; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 167 GCCAAATTGATCCCTCTTCAGACATATGACAGACACACCTTCAGATATTTAAG 226  
DB 58 CCCAAATTGATCCCTCTTCAGACATATGACAGACACACCTTCAGATATTTAAG 117  
QY 227 AGCTTCAACAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTG 286  
DB 118 AGCTTCAACAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTC 177  
QY 287 CGGCTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGTATTTGCTCAGACTTTA 346  
DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGAAATGAAGTTGTATTTGCTCAGACTTTA 237  
QY 347 CACATAGAGAGTTCACACCTGGCTCCGCC-AATCCCGACAAACAGTTCCTCATCTCCCT 405  
DB 238 CACATAGAGAGTTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTCTGATCTCCCT 297  
QY 406 CCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTAC 465  
DB 298 CCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTAC 357  
QY 466 GATCTTTTATATGCGATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGACGCG 525  
DB 358 GATCTTTTATATGCGATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGACGCG 417  
QY 526 ACAGACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGCACCAAGAGATGAGGAG 585  
DB 418 ACTGACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGCATGATGAGAGAGGAGGAA 477  
QY 586 GAAGAGGAA 594  
DB 478 GAAGAGGAA 486

RESULT 9  
AAX01282  
ID AAX01282 standard; cDNA to mRNA; 2174 BP.  
AC  
XX AAX01282;  
DT 09-APR-1999 (first entry)  
XX  
DE Human DSCR1 coding sequence.

XX DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;  
KW Central Nervous System development; mental retardation; heart defect; ds.  
XX Homo sapiens.

Key Location/Qualifiers  
CDS 49..564  
FT /\*tag= a  
FT polyA\_signal 1541..1546  
FT /\*tag= b  
FT polyA\_signal 2132..2137  
FT /\*tag= c

XX US5869318-A.  
XX  
XX 09-FEB-1999.  
XX  
XX 07-JUN-1996; 96US-0665040.

XX 07-JUN-1995; 95ES-0001140.  
PR (PALL/) PALLEJA X E.  
XX  
XX Fuentes JJ, Palleja XE, Pritchard M;  
PI WPI: 1999-152781/13.  
DR P-PSDB; AAW73898.  
XX  
PT DNA encoding foetal brain proteins - believed to be associated with  
PT Down's syndrome  
XX  
PS Claim 1; Column 15-18; 19pp; English.  
XX  
CC This sequence is the Down's Syndrome critical region 1 (DSCR1) gene  
CC of the invention. The DSCR1 gene was found to be located in the  
CC q22.1-22.2 region of human chromosome 21. An increase in the transient  
CC expression of DSCR1 mRNA in the brains of young rats, compared to  
CC expression levels in the brains of adult rats, suggests an important role  
CC for DSCR1 during the development of the Central Nervous System (CNS), and  
CC that over expression of DSCR1 may be involved in pathogenic abnormalities  
CC of mental retardation and/or heart defects as found in Down's syndrome  
CC patients.  
XX  
SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;

Query Match 56.8%; Score 340.2; DB 20; Length 2174;  
Best Local Similarity 88.6%; Pred. No. 1.5e-99;  
Matches 380; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 167 GCCAAATTGATCCCTCTTCAGACATATGACAGGACACACCTTCAGTATTTAAG 226  
DB 58 GCCAAATTGATCCCTCTTCAGGACGTATGACAGGACATCACCCTTCAGTATTTAAG 117  
QY 227 AGCTTCAACAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTG 286  
DB 118 AGCTTCAACAGTGTCCGGATAAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177  
QY 287 CGGCTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGTATTTGCTCAGACTTTA 346  
DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGAAATGAAGTTGTATTTGCTCAGACTTTA 237  
QY 347 CACATAGAGAGTTCACACCTGGCTCCGCC-AATCCCGACAAACAGTGCATATAAATTAC 405  
DB 238 CACATAGAGAGTTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTCTGATCTCCCT 297  
QY 406 CCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTAC 465  
DB 298 CCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTAC 357  
QY 466 GATCTTTTATATGCGATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGACGCG 525  
DB 358 GATCTTTTATATGCGATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGACGCG 417  
QY 526 ACAGACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGCACCAAGAGATGAGGAG 585  
DB 418 ACTGACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGCATGATGAGAGAGGAGGAA 477  
QY 586 GAAGAGGAA 594  
DB 478 GAAGAGGAA 486

RESULT 10  
AAX01774  
ID AAX01774 standard; cDNA; 442 BP.  
XX  
XX AAC01774;  
XX AC  
XX AC  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 1772.









QY 220 TTTAGAGCTTCAACGTCGCGATTAACCTTACGACCAACCCCTTATCTGACGCGATGC 279  
 DB 330 ATTGAGGCTTTCAGACGCTGTCGATTAACCTTACGACCAATCTGACGCGCGAGC 389  
 QY 280 CAGGCTCGGCTGCAACAGCCGAGTCTCTGGGAGGAAATGAAGTTGATTTTGTCTCA 339  
 DB 390 TAGGATAGAGCTTCTATGAACCCATTCAGAGGAGGAAAAATTAAGCTCTACTTGTGACA 449  
 QY 340 GACT-----TTACACATAGGAAGTTACACCTGGCTCC-GCCAAATCCCGACA 386  
 DB 450 GGTTCAGACTCCAGAGACAGATGAGAGCAAACTGCACTTGGCTCCACCCCGCTGCCAA 509  
 QY 387 ACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGCTGCGTGGGAAACAGTAGAAGATGC 446  
 DB 510 ACAGTTTCTCATCTGCGCCCTCTCTCCCACTTGTAGCTGGAGCCCATCATGATGC 569  
 QY 447 CACCCCGCTCAATTAATGATCTTTATATGCACTTCCAAAGCTGGGGCCAGAGAGAA 506  
 DB 570 CACGCGAGTCCCTCACTATGACCTCCTATGCTGTGGCCAAACTAGAGCCAGAGAGAA 629  
 QY 507 GTATGACTGCTGCAAGCAGACACCACTCCAGTGTGGTGGTCCAGTGTGAGAG 566  
 DB 630 GTATGAGCTCCATGCAAGGAGCTGAGTCCACCCCAAGTGTGTGACAGTGTGAGAG 689  
 QY 567 TGACCAAGAGATGAGAGAGA 587  
 DB 690 TGACATAGAGAGAGAGAGA 710  
 RESULT 13  
 ABL65187  
 ID ABL65187 standard; DNA; 3184 BP.  
 AC ABL65187;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Lung cancer related gene sequence SEQ ID NO:3524.  
 XX  
 KM Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KM cytotoxic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
 KM gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 XX 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-23133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234044P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR

PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 28-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-23712P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 PR  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI, 2002-188264/24.  
 XX  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 PT  
 PS Claim 1; SEQ ID 3524; 44p; English.  
 XX  
 XX

CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (i)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL65187  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (i) has cytotoxic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 CC  
 XX  
 XX Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other:

Query Match 33.8%; Score 202.2; DB 24; Length 3184;  
 Best Local Similarity 65.9%; Pred. No. 1.1e-54;  
 Matches 330; Conservative 0; Mismatches 158; Indels 13; Gaps 2;

QY 100 CTACAACTTTAGCTCCGATGCTTGTGTGGCAACGATGATGCTTCAGGAAAGTGA 159  
 DB 210 CTGTGATGTTCCACCTGCTGTGCTGTGTGTGGATGTCGAGCTTTTACCAATCAGGA 269  
 QY 160 GACCAAGGCCAAATTGAATCCCTTTCAGAACATATGACAGACACCACTTCAGTA 219  
 DB 270 GGTTAAGGAAAAATTTGGGGAGCTGTTTCGACTATATGATGACTGTGACCTTCAGCT 329  
 QY 220 TTTAAGACTTCAACGTCGCGATTAACCTTACGACCAACCCCTTATCTGACGCGATGC 279  
 DB 330 ATTGAGGCTTTCAGACGCTGTCGATTAACCTTACGACCAATCTGACGCGCGAGC 389  
 QY 280 CAGGCTCGGCTGCAACAGCCGAGTCTCTGGGAGGAAATGAAGTTGATTTTGTCTCA 339

Db 390 TAGATAGAGTTTCATGAACCCCAATTCAGAGGGAATAATTAAGCTCTACTTTGCACA 449  
 QY 340 GACT-----TTACACATAGGAAGTTTCACACTGGCTCC-GCCAAATCCCGACAA 386  
 Db 450 GGTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTTGGCTCCACCCAGCTGCCAA 509  
 QY 387 ACAGTTCTCATCTCCCTCCGAGCTCTCTCCGGTTGGCTGGAAACAAGTAGAAGATGC 446  
 Db 510 ACAGTTTCTCATCTCGGCCCTTCTCTCCCACTGTTAGCTGCGAGCCCACTCAACGATGC 569  
 QY 447 CACCCCGCTCATAAATACCATCTTTTATATGTCATCTCCAAAGCTGGGGCCAGGAGAA 506  
 Db 570 CAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACACAGAGAGAA 629  
 QY 507 GTATGAATGCTGACGCGACAGACCACTCCCAAGTGGTGGTCCAGTGTGTGAGAG 566  
 Db 630 GTATGAGCTCATGCGAGGACTGAGTCCACCCCAAGTGTGCTGTCACGTGTGCGACAG 689  
 QY 567 TGACCAAGAGAAATGAGGAGCA 587  
 Db 690 TGACATAGGAAGAGAGGA 710

## RESULT 14

AAF25338

ID AAF25338 standard; cDNA; 934 BP.

XX AAF25338;

XX 30-APR-2001 (first entry)

XX Nucleotide sequence of a human detoxification protein.

XX Human; detoxification protein; DETX; cancer; leukaemia; melanoma;

XX adenocarcinoma; autoimmune disorder; inflammatory disorder;

XX rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;

XX psoriasis; ulcerative colitis; infection; cell proliferative disorder;

XX actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX 20..724

XX /tag= a

XX /product= "detoxification protein"

XX 20..100

XX /tag= b

XX sig\_peptide

XX WO200104305-A2.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000WO-US18509.

XX 07-JUL-1999; 99US-0142678.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H;

XX WPI; 2001-147193/15.

XX P-PSDB; AAB31788.

XX New human detoxification protein and polynucleotide, useful for

XX diagnosis, prevention and treatment of autoimmune/inflammatory

XX disorders and cell proliferative disorders including cancer.

XX Claim 5; Page 79; 79pp; English.

XX PS

XX The present sequence encodes a human detoxification polypeptide (DETX).

XX DETX and its (ant)agonists are useful for preventing or treating

XX disorders associated with decreased or increased expression or activity

XX of DETX. DETX polypeptides are useful for screening compounds that

CC specifically binds to DETX and for identifying (ant)agonists.  
 CC Diseases prevented, treated and diagnosed include cancers (e.g.  
 CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,  
 CC breast, kidney, liver, pancreas, prostate and uterus),  
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,  
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative  
 CC colitis), bacterial, fungal, parasitic infections and cell  
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,  
 CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as  
 CC antagonists, as a targeting or delivery mechanism for bringing  
 CC pharmaceutical agents into contact with cells or tissues expressing  
 CC DETX and for diagnosis of DETX-related disorders.

XX Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;

Query Match 33.0%; Score 197.6; DB 22; Length 934;

Best Local Similarity 64.6%; Pred. No. 1.8e-53;

Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;

QY 87 TTAGGGACTTTAGTACAAATTTTAGCTCCCTGATTGCTGTGGCAACGATGATGCT 146  
 Db 138 TCACTGACTTCAATGACCTCCCAACTCGTTGTTGCGTGCATGTTCCACGATCAGTGT 197  
 QY 147 TCAGCGAAAGTGAGACCCAGGCGCAAAATTTGAATCCCTCTTCAGAAATATGACAAAGGACA 206  
 Db 198 TTGAAGGAGAGAGAGAGCAAGGAAAAATTTGAGGGACTGTTCCGACTTATGATGACTGTG 257  
 QY 207 CCACCTTCCAGTATTTTAAGAGCTTCAACGTTGCGGATAACTTCAGCAACCCCTTAT 266  
 Db 258 TGACGTTCCAGCTATTTAAGAGTTTCAGACGTTGCGGTATAAACTTCAGCAATCCCTAAAT 317  
 QY 267 CTGACGCCGATGCCAGGTGCGGCTGCACAGACCGAGTTCTCGGGAGGAAATGAAGT 326  
 Db 318 CTGACGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGAAAAATTAAGC 377  
 QY 327 TGTATTTGCTCAGACT-----TTACACATAGGAAGTTTCACACTGGGTCC-G 373  
 Db 378 TCTACTTTGCACAGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTGGCTCCAC 437  
 QY 374 CCAATCCCGACAAACAGTTTCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGTGAAC 433  
 Db 438 CCCAGCTTCCCAACAGTTTCTCATCTCGCCCTTCTCTCCCACTGTGTGGCTGGCAGC 497  
 QY 434 AAGTAGAAGATGCCACCCCTGCATAAATTCAGATCTTTTATATGCCATCTCCAAAGCTGG 493  
 Db 498 CCATCAACGATGCCACCGCAGTCCCTCACTATGACTCTCTATGCTGTGGCCAACTAG 557  
 QY 494 GGCAGGAGAGAGTATGAATGCAATGCAAGCAGACAGACCACTCCCACTGTGTGGGTCC 553  
 Db 558 GACCAGGAGAGAGTATGAGTCCATGCAAGGAGTGAAGTCCACCCCAAGTGTGCTGCTGC 617  
 QY 554 ACGTGTGTGAGAGTGACCAAGAGATGAGGAGGA 587  
 Db 618 ACGTGTGCACAGTGACATAGAGGAAGAAGAGGA 651

## RESULT 15

ABA91457

ID ABA91457 standard; cDNA; 3159 BP.

XX ABA91457;

XX 18-APR-2002 (first entry)

XX Human Down syndrome critical region 1-like 1 protein cDNA.

XX Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;

XX Down syndrome; Alzheimer's disease; dementia; transgenic;

XX neuroprotective; nontoxic; anticonvulsant; diagnosis;

XX gene therapy; gene; ss.

XX Homo sapiens.

PH Key Location/Qualifiers  
 FT CDS 185..952  
 FT /tag="a  
 FT /product="DSCR1L1"  
 FT /trans\_except="(pos:215..217, aa:Xaa)  
 FT /note="Xaa = unknown"  
 PN MO200204513-A2.  
 XX  
 PD 17-JAN-2002.  
 PF 11-JUL-2001; 2001WO-US21982.  
 PR 11-JUL-2000; 2000US-0614474.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 PI Loring JF, Tingley DM, Edwards CM, Streeter DG;  
 DR WPI; 2002-164633/21.  
 DR P-PSDB; AAM50760.  
 PT Novel Down syndrome critical region 1-like 1 protein and nucleic acid  
 PT encoding the protein useful for diagnosis and treatment of Alzheimer's  
 PT disease, Down syndrome and other forms of dementia  
 XX  
 PS Claim 2; Page 46-47; 54pp; English.  
 CC The present sequence is that of cDNA clone Incyte ID No: 247500.5  
 CC encoding novel Down syndrome critical region 1-like 1 protein  
 CC (DSCR1L1 alpha, see AAM50760). Northern analysis indicated  
 CC expression of DSCR1L1 alpha in various libraries, with the highest  
 CC abundance in tissues from the nervous system, including tissues  
 CC associated with schizophrenia, Huntington's disease, epilepsy and  
 CC amyotrophic lateral sclerosis. An absence of DSCR1L1 expression  
 CC was observed in 7 of 8 libraries from subjects with Alzheimer's  
 CC disease. A claimed method for detecting differential expression of  
 CC a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down  
 CC syndrome, Alzheimer's disease and other forms of dementia. A probe  
 CC from such a nucleic acid is useful for identifying naturally  
 CC occurring molecules encoding DSCR1L1 alpha, allelic variants or  
 CC other molecules. The nucleic acids are useful for producing  
 CC transgenic cell lines or organisms which model human disorders.  
 CC They may also be used in gene therapy, and to screen for ligands  
 CC which specifically bind the nucleic acid molecule, such as  
 CC a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,  
 CC peptide, transcription factor, repressor or regulatory molecule,  
 CC for use as a therapeutic.  
 XX  
 SQ Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;  
 Query Match 33.0%; Score 197.6; DB 24; Length 3159;  
 Best Local Similarity 64.8%; Pred. No. 3.4e-53;  
 Matches 312; Conservative 0; Mismatches 169; Indels 13; Gaps 2;  
 Oy 87 TTAGGACCTTACCTACATTTAGCTCCCTGATGCTGTGCAAAACGATGATCTT 146  
 Db 366 TCACCTACTTCAATGACCTCCCACTCCTGTTGGTGCAATGTTCAACAGTCAGTGT 425  
 Oy 147 TCAGCGAAGTGAAGACCAAGGCAAAATTTGAATCCCTCTTCAGAACATATGACAGGACA 206  
 Db 426 TTGAAGGAGAAAGAGCAAGAAATTTGAGGAGACTGTTCCGACTTATGATGACTGTG 485  
 Oy 207 CCACCTTCCAGTATTTTAAGAGCTTCAAAAGCTGCCGATTAACCTTCAGCAACCCCTTAT 266  
 Db 486 TGAAGTTCAGCTATTTAAGAGTTTCAGAGCTGTCGTATTAACCTTCAGCAATCTTAAT 545  
 Oy 267 CTGAGGCGATGCGAGGCTGCGGCTGCACAAAGCGAGTCTCGGAGAGAAATGAAAGT 326  
 Db 546 CTGAGGCGCGAGGCTGAGATGAGCTTCATGAAACCAATTCAGAGGAGAAATTAAGC 605  
 Oy 327 TGTATTTTGTCTCAGACT-----TTACATAGGAAGTTCAACACCTGGCTCC-G 373

Db 606 TCTACTTTGCACAGGTTCAAGCTCCAGAGACAGATGAGAGCAAACTGCACTTGGCTCCAC 665  
 Oy 374 CCAATCCCGACAAACAGTTCTCATCTTCCCTCCGAGCTCTCTCCGTTGGCTGAAAC 433  
 Db 666 CCCAGCTGCGCAAAACAGTTCTCATCTGCGCCCTCTCTCCCACTGTTGGCTGGGAGC 725  
 Oy 434 AAGTAGAAGATGCCACCCCGTCAATAATTAGATCTTTTATATGCAATGCAAGCTGG 493  
 Db 726 CCATCAAGATGCGACGCACTCTCACTATGACCTCTCTATGCTGGCCAAACTAG 785  
 Oy 494 GAGCAGAGAGAAATGAACTGATGATGATGATGATGATGATGATGATGATGATGATG 553  
 Db 786 GACCAAGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845  
 Oy 554 ACGTGTGTGAGATGACCAAGAGAAATGAGAGGA 587  
 Db 846 ACGTGTGTGAGATGACCAAGAGAAATGAGAGGA 879  
 Search completed: December 14, 2002, 15:18:56  
 Job time : 103.824 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 18:11:26 ; Search time 139.234 Seconds  
(without alignments)  
3202.482 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFRSYNPSLIACVAND.....RPKPKIOTRRPYTPHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09782953/runat\_11122002\_114428\_17330/app\_query.fasta\_1.1173  
-DB=N Geneseq 101002 -QWTF=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 @CNG\_1\_1\_220 @runat\_11122002\_114428\_17330 -NCPU=6 -ICPU=3  
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq 101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1043	100.0	597	24	AAD30153	Mouse MCIP associa
2	970.5	93.0	2331	24	AAD30155	Human MCIP associa
3	957.5	91.8	2348	24	AAS94805	Human DNA sequence
4	920.5	88.3	597	24	AAD30152	Human MCIP associa
5	875	83.9	2212	24	AAD30156	Human MCIP associa
6	868	83.2	599	24	AAD30151	Human MCIP associa
7	847.5	81.3	2173	24	AEN96843	Gene #3341 used to
8	844.5	81.0	2174	20	AAK01282	Human DSCR1 coding
9	830	79.6	2358	24	AAF18328	Lung cancer associ
10	632.5	60.6	594	24	AAD30154	Mouse MCIP associa
11	623	59.7	934	22	AAF25338	Nucleotide sequenc
12	623	59.7	3159	24	ABA91457	Human Down syndrom
13	615.5	59.0	3184	24	AAD30157	Human MCIP associa
14	615.5	59.0	3184	24	ABL61768	Colon adenocarcino
15	615.5	59.0	3184	24	ABL65187	Lung cancer relate
16	592.5	56.8	828	24	AAD30158	Human MCIP associa
17	591.5	56.7	720	24	AAD30159	Human secreted pro
18	537	51.5	442	21	AAC01774	Rat Down syndrome
19	536.5	51.4	615	24	ABA91463	Rat Down syndrome
20	528.5	50.7	1021	24	ABA91458	Drosophila melanog
21	372	35.7	1820	23	ABL29301	Human foetal liver
22	372	35.7	12550	23	ABL29300	Human brain expres
23	285.5	27.4	412	22	ABA56172	Human foetal liver
24	285.5	27.4	412	22	AAK04364	Human brain expres
25	285.5	27.4	412	22	AAI14449	Probe #4382 for ge
26	285.5	27.4	412	22	AAI35821	Probe #4507 used t
27	285.5	27.4	412	22	AAI04272	Probe #4263 used t
28	285.5	27.4	412	22	ABSO4419	Human genome-deriv
29	285.5	27.4	446	22	ABA43406	Human breast cell
30	285.5	27.4	446	22	ABA33852	Human foetal liver
31	285.5	27.4	446	22	ABA23598	Probe #2064 for ge
32	285.5	27.4	446	22	AAK02113	Human brain expres
33	285.5	27.4	446	22	AAK27563	Human bone marrow
34	285.5	27.4	446	22	AAI12148	Probe #2081 for ge
35	285.5	27.4	446	22	AAI33493	Probe #2179 used t
36	285.5	27.4	446	22	AAI02061	Probe #2052 used t
37	285.5	27.4	446	24	ABSO2041	Human genome-deriv
38	272	26.1	640	24	ABO59719	Human colon cancer
39	266.5	25.6	486	22	ABA42132	Human breast cell
40	266.5	25.6	486	22	ABA25254	Human foetal liver
41	266.5	25.6	486	22	ABA22343	Probe #809 for gen
42	266.5	25.6	486	22	AAK00816	Human brain expres
43	266.5	25.6	486	22	AAK26270	Human bone marrow
44	266.5	25.6	486	22	AAI10903	Probe #836 for gen
45	266.5	25.6	486	22	AAI32163	Probe #849 used to

ALIGNMENTS

RESULT 1

AAD30153

ID AAD30153 standard; DNA; 597 BP.

XX AAD30153;

XX 17-MAY-2002 (first entry)

XX Mouse MCIP associated DNA #3.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

XX heart failure; cardiomyopathy; heart disease; mouse; ds.

XX Mus musculus.

XX Key

XX Location/Qualifiers

XX 1..597

XX /\*tag= a

FT /product= "Mouse MCIP associated protein #3"

XX MO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001MO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI, 2002-179698/23.

XX P-Psdb; AAE18912.

XX Screening for modulators of muscle calcineurin interacting protein

XX (MCIP) binding, expression or phosphorylation, useful for treating

XX cardiac hypertrophy or heart failure, comprises mixing MCIP,

XX calcineurin and a test compound -

XX Disclosure; Page 150-151; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or

XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin

XX may be used for treating cardiac hypertrophy and heart failure.

XX Antibodies to MCIP can be used in characterizing the MCIP content of

XX healthy and diseased tissues and subsequently for determining the

XX presence or absence of cardiomyopathy or as predictor of heart disease.

XX The present sequence is mouse MCIP associated DNA.

XX Note: This sequence has been described as human MCIP3 encoding DNA

XX in the specification, however the sequence seems to be a polynucleotide

XX encoding a MCIP associated protein.

XX SQ Sequence 597 BP; 170 A; 156 C; 142 G; 129 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 4.5e-109 Length: 597

XX Score: 1043.00 Matches: 198

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 100.00% Indels: 0

XX DB: 24 Gaps: 0

XX US-09-782-953-7 (1-198) x AAD30153 (1-597)

XX 1 MetAspPheArgAspPheSerTyrAsnPheserSerleuileAcyValAlaAsnAsp 20

XX 1 ATGGATTTTGGGACTTACATCAATTTTAACTCCCTGATGCTGCGCAAAAGAT 60

XX 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerleuPheArgTrrTyrAsp 40

XX 61 GATGCTTTCAGCAAGTGAAGCCAGGCGCAAAATTTGAAATCCCTCTTCAGAAACATATGAC 120

XX 41 LysAspPheThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheserAsn 60

XX 121 AAGGACCCACCTCCAGTATTTTAAAGCTTCAACCTGTCGAGTAACTTCAGCAAC 180

XX 61 ProleuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

XX 181 CCTTATCTCGACGCGGATGCCAGCTCGGCTGCAAAAGACCGAGTCTCGGGGAAGGAA 240

XX 81 MetLeuLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100

DB 241 ATGAAGTTGATTTTGGCTCAGACTTACACATAGGAAGTTCAACCTGGCTCCGCCAAT 300

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGlnVal 120

DB 301 CCCGACAAACAGCTTCTTCATCTCCCTCCGCTCTCTCTCCCTGGCTGGAAACAGTA 360

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140

DB 361 GAGATGCCACCCCGCTCATATAATTAACATCTTTATATCCCATCTCCAACTGGGGCCA 420

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160

DB 421 GGAGAGAGATGATGACATGACGACGACGACCCCACTCCAGTGTGTGCTCAGCTG 480

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 481 TGTGAGACTGACCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198

DB 541 AAGCCAAATCATCCAGACACGAGACCGAGTACACCGATCCACCTTAGC 594

RESULT 2

AAD30155

ID AAD30155 standard; DNA; 2331 BP.

XX AAD30155;

XX AC

XX 17-MAY-2002 (first entry)

XX DT

XX Human MCIP associated DNA #1.

XX DB

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

XX heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX OS

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT 144..797

XX FT CDS /\*tag= a

XX FT /product= "Human MCIP associated protein #1"

XX PN MO200204491-A2.

XX PD 17-JAN-2002.

XX PF 06-JUL-2001; 2001MO-US21662.

XX PR 07-JUL-2000; 2000US-216601P.

XX PR 13-FEB-2001; 2001US-0782953.

XX PA (TEXA ) UNIV TEXAS SYSTEM.

XX PA (WILL/) WILLIAMS S R.

XX PA (ROTH/) ROTHERMEL B.

XX PI Williams SR, Rothermel B;

XX WPI, 2002-179698/23.

XX DR P-Psdb; AAE18914.

XX PT Screening for modulators of muscle calcineurin interacting protein

XX (MCIP) binding, expression or phosphorylation, useful for treating

XX cardiac hypertrophy or heart failure, comprises mixing MCIP,

XX calcineurin and a test compound -

XX PS Claim 72; Page 155-157; 174pp; English.

XX CC The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or



Db 451 TCCAGACAGCAGTTTCGATCTCCCTCCCTCCGCGAGTGGAACAAGT 510  
 Oy 120 TGTAAATAThProValIleAsnTyraPheLeuTyraIleSerLysLeuGlyPr 140  
 Db 511 GGAAGATGGACCCGATCACTAATATGATCTTTATATGCACTCCAGCTGGGCC 570  
 Oy 140 OGlyGlyuTyrgLjLeuHlaIaIaThraPProthProSerValIaIaHlaIa 160  
 Db 571 AGGGAAAGTATGATTCACGACGCGACTGACACCACTCCGACGTCGTGTCATGT 630  
 Oy 160 TCGGtLserAaPglngLlAaIngLjGluGluGluGluGluGluGluGluGluGlu 180  
 Db 631 ATGTGAGTGTATCAAGAG--AAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 687  
 Oy 180 GlysPProLysIleIleGlnThraArgArpProGluTyrrThrProIleHlaLeuSer 198  
 Db 688 TAAAGCAAAATTTTCAACACGAGAGCCGAGATACAGCCGATCCACTCAGC 742  
 RESULT 4  
 AAD30152  
 ID AAD30152 standard; DNA; 597 BP.  
 AC AAD30152;  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated DNA #2.  
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..597 a  
 CDS /\*tag= a /product= "Human MCIP associated protein #2"  
 PN MO20204491-A2.  
 PD 17-JAN-2002.  
 PF 06-JUL-2001; 2001WC-US21662.  
 PR 07-JUL-2000; 2000US-21661P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 PI Williams SR, Rothermel B;  
 DR WPI, 2002-179698/23.  
 DR P-PSDB; AAE18911.  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 PS  
 XX Claim 95; Page 147-148; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the

CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 XX  
 SQ Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 3.69e-95 Length: 597  
 Score: 920.50 Matches: 178  
 Percent Similarity: 92.42% Conservative: 5  
 Best Local Similarity: 89.90% Mismatches: 12  
 Query Match: 88.26% Indels: 3  
 DB: 24 Gaps: 1  
 US-09-782-953-7 (1-198) x AAD30152 (1-597)  
 Oy 1 MetAaPheArGaPheSerTyraPheSerLeuIleAlaCyValAlaAsnAaP 20  
 Db 10 GTGATCTGCAGACCTGCGG-----AGGCCACACATCGCTCCGACCTGACCGG 60  
 Oy 21 AaPValPheSerGluSerGluThraAlaLysPheGluSerLeuPheArGThrTyraP 40  
 Db 61 CGCGTTCGTGAGACGGCCCTGTCGGGCCAAATTGCAATCCCTTCAGAAATATGAC 120  
 Oy 41 LysAaPThrThrPheGlnTyrrPheLysSerPheLysArGValArgIleAaPheSerAaP 60  
 Db 121 AAGGACACACCTTCAGATATTTAAGAGCTTCAAGCGTGCAGATTAACCTTCAGGAA 180  
 Oy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHlaIaIaGlnPheLeuGlyLysGlu 80  
 Db 181 CCTTATTTGACGCGGATGCCAGGTGGCGGTGCACAAAGCGAGTTCCTGGGAAAGGAA 240  
 Oy 81 MetLysLeuTyrrPheAlaGlnThreuhIleIleGlySerSerHlaLeuAlaPProAaP 100  
 Db 241 ATGAAGTGTATTGCTCAGACTTTACATAGAACTTCACACTCGGCTCCGCCCAAT 300  
 Oy 101 ProAaPLeuGlnPheLeuIleSerProProAlaSerProProValGlyTyrrLysGlnVal 120  
 Db 301 CCGGACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGCTGGGAAACAAAGTA 360  
 Oy 121 GluAaPAlaThrProValIleAsnTyraPheLeuTyraIleSerLysLeuGlyPro 140  
 Db 361 GAAAGATGCCACCCCGCTCATTAATAGATCTTTATATGCACTCCAGACTGGGCCA 420  
 Oy 141 GlyGlyuTyrgLjLeuHlaIaIaThraPProthProSerValIaIaHlaIa 160  
 Db 421 GGAGAGAAGTATGAACCTGATGCAGCGACAGACCCCACTCCAGTGTGTGTCACCTG 480  
 Oy 161 CysGluSerAaPglngLlAaIngLjGluGluGluGluGluGluGluGluGluGlu 180  
 Db 481 TGTGAGACTGACCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACC 540  
 Oy 181 LysPProLysIleIleGlnThraArgArpProGluTyrrThrProIleHlaLeuSer 198  
 Db 541 AAGCCAAATATCATCCAGACGAGAGCCGAGTACACCGATCCACTCAGC 594  
 RESULT 5  
 AAD30156  
 ID AAD30156 standard; DNA; 2212 BP.  
 AC AAD30156;  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated DNA #2.  
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 25..618 a  
 CDS /\*tag= a



```

PT XX /product= "Human MCIP associated protein #2"
PN WO200204491-A2.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-US21662.
XX 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX (TEXA ) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
PA (ROTH/) ROTHERMEL B.
XX Williams SR, Rothermel B;
XX WPI; 2002-179698/23.
DR P-PSDB; AAE18915.
XX
XX Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
XX Example 1; Page 159-161; 174pp; English.
XX
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC transcription of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterising the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is human MCIP associated DNA.
CC Note: This sequence has been described as murine MCIP splice variant in
CC the specification, however the sequence seems to be a polynucleotide
CC encoding a MCIP associated protein.
XX
XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 3,31e-89 Length: 2212
Score: 875.00 Matches: 171
Percent Similarity: 90.40% Conservative: 8
Best Local Similarity: 86.36% Mismatches: 15
Query Match: 83.89% Indels: 4
DB: 24 Gaps: 2

```

US-09-782-953-7 (1-198) x AAD30156 (1-2212)

```

Qy 1 MetAspPheArgAspPheSerTyArgAspPheSerLeuIleAlaCysValAlaAsnAsp 20
Db 34 GTGACCTGAGACCTGCCCC-----AGCGCCACCATCGCTGTACCTGGACCCG 84
Qy 21 AspValPheSerGluSerGluThrArgAlaValPheSerLeuPheArgThrTyArg 40
Db 85 CGCGTGTCTGTGACCGCCCTGTCGGGCCCAATTTAGTCCCTCTTTAGGAGGTATGAC 144
Qy 41 LysAspThrThrPheGlnTyPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACCCTTCAGTATTATTAAGAGCTTCAACGAGTCCAGATAAATTCACGAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGACGACGATGCGAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 264
Qy 81 MetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

```

```

Db 265 ATGAAGTTATATTTTGTCTCAGACCTTTACATAGGAAGCTCACACCTGGCTCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
Db 325 CCAGACAGCAGTTTCTGATCTCCCTCCGCTCTCCGCCAGTGGATGGAAACAAGTG 384
Qy 121 GluAspAlaThrProValIleAsnTyArgLeuLeuTyArgAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCAGTCATAAATATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 444
Qy 141 GlyGluLysTyGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAAGTATGAATTGCACGCGACTGACACCACTCCCGAGCTGGTGGTCCATGTA 504
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTGTATCAGAG--AAGGAGGAAGAGAGAAATGGAAGAATAGGAGACCT 561
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
Db 562 AAGCCAAAATATTCAGACACGAGGAGCGGAGTACACGCCGATCCACCTCAGC 615
RESULT 6
AAD30151
ID AAD30151 standard; DNA; 599 BP.
XX
AC AAD30151;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated DNA #1.
XX
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; human; ds.
XX
OS Homo sapiens.
XX
PN WO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
PA (ROTH/) ROTHERMEL B.
PI Williams SR, Rothermel B;
XX
DR WPI; 2002-179698/23.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Disclosure; Page 147; 174pp; English.
XX
CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC transcription of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterising the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is human MCIP associated DNA.

```

XX SQ Sequence 599 BP; 164 A; 150 C; 149 G; 136 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 3.4e-89 Length: 599  
 Score: 868.00 Matches: 170  
 Percent Similarity: 98.84% Conservative: 0  
 Best Local Similarity: 98.84% Mismatches: 2  
 Query Match: 83.22% Indels: 1  
 DB: 24 Gaps: 0

US-09-782-953-7 (1-198) x AAD30151 (1-599)

OY 2 AAPPHEARAPAPHESErTYrAsnPheSeSeLeuLeaLcYvAlaLlAsnAsp 21  
 DB 83 GATTTTAGGACTTTAGCTACAAATTTTACTCCGATTGCTTGTGCAACATGAT 142  
 OY 22 ValPheSeGluSeGluThrAlaGAlaLysPheGluSeLeuPheArGThrTYrAspLys 41  
 DB 143 GTCTTCAGCGAAAGTGAGACAGGCGCAATTTGAAATCCCTTCAGAACATATGACAG 202  
 OY 42 AAPPThrPhcGluTYrPheLysSePheLysArgValArgIlleAsnPheSeAsnPro 61  
 DB 203 GACACACCTTCAGATATTTAAGAGCTTCAACGTCGATTAACCTTCAGCAACCCC 262  
 OY 62 LeuSeAlAlAspAlaArgLeuArgLeuHlAlYsThrGluPheLugLlYsGluMet 81  
 DB 263 TTATCTGACAGCGAGTCCAGGCTGGCGTGCACCAAGCCGAGTCTGGGAAAGAAATG 322  
 OY 82 LysLeuTYrPheAlAGlnThrLeuHlAlleGlySeSeThlLeuAlaProProAsnPro 101  
 DB 323 AAGTGTATTGTGCTCAGACTTACACATAGAGAGTTCCACCTGCTCCGCAT-CCC 381  
 OY 102 AAPPYsGluPheLeuLlSeSeProProAlaSeProProValGlyTYrPheGlnValGlu 121  
 DB 382 GACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGGCTGGGAAACAGTAGAA 441  
 OY 122 AAPPAlaThrProValIlleAsnTYrAspLeuLeuTYrAlIlleSeTyAlaLugLlYsProGly 141  
 DB 442 GATGCCACCCCTCATTAATTAGATCTTTATATGATCCATCCCAAGCTGGGGCCAGGA 501  
 OY 142 GluLYsTYrGluLeuHlAlaAlaThrAspProThProSeValValValHlSvalCYs 161  
 DB 502 GAGAAATATGATCATGATGACGACAGACACACCTCCAGTGTGTGTGTCACGTGTG 561  
 OY 162 GluSeAspGlnGluAsnGluGluGluGluGlu 173  
 DB 562 GAGAGTGACCAAGAGATGAGAGAAAGAGAGAG 597

RESULT 7  
 ABN96843  
 ID ABN96843 standard; DNA; 2173 BP.  
 XX AC ABN96843;  
 XX DT 13-AUG-2002 (first entry)  
 XX DE Gene #3341 used to diagnose liver cancer.  
 XX KW Gene; liver cancer; db; hepatocellular carcinoma; hepatotropic;  
 KW metastatic liver tumor; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX OS Homo sapiens.  
 XX PN WO200229103-A2.  
 XX PD 11-APR-2002.  
 XX PF 02-OCT-2001; 2001MO-US30589.  
 XX PR 02-OCT-2000; 2000US-237054P.  
 XX

PA (GENE-) GENE LOCIC INC.  
 XX XX  
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX WPI; 2002-426119/45.  
 DR  
 XX  
 XX  
 FT Diagnosing and detecting the progression of liver cancer,  
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
 PT involves detecting the level of expression of two or more genes in a  
 PT liver tissue sample -  
 XX  
 PS Claim 1; SEQ ID NO 3341; 298pp; English.

XX XX  
 CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumour in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 4.29e-86 Length: 2173  
 Score: 847.50 Matches: 162  
 Percent Similarity: 97.63% Conservative: 3  
 Best Local Similarity: 95.86% Mismatches: 3  
 Query Match: 81.26% Indels: 1  
 DB: 24 Gaps: 1

US-09-782-953-7 (1-198) x ABN96843 (1-2173)

OY 30 AlalySePheGluSeLeuPheArGThrTYrAspLysAAPPThrPhcGlnTYrPheLys 49  
 DB 58 GCCAAATTTAGTCCCTCTTAAGACGTATACAAAGACATCCTTCAGATTTTAAG 117  
 OY 50 SerPheLysArgValArgIlleAsnPheSeAsnProLeuSeAlAlaAspAlaArgLeu 69  
 DB 118 AGCTTCAACAGATCAGATTAATTAACCTTCAGCAACCCCTTCGAGAGAGTCCAGGCTC 177  
 OY 70 ArgLeuHlAlYsThrGluPheLugLlYsGluMetLysLeuTYrPheAlAGlnThrLeu 89  
 DB 178 CAGCTGACATTAAGACTGAGATTTCTGGAAAGAAATGAAATATATTGTGCTCAGACCTTA 237  
 OY 90 HlIlleGlySeSeThlLeuAlaProProAsnProAspLysGlnPheLeuLlSeSePro 109  
 DB 238 CACATAGAGAGCTCACCTGCTCGCCAAATTCAGACAGACATTTCTGATCTCCCT 297  
 OY 110 ProAlaSeProProValGlyTYrPheGlnValGluAspAlaThrProValIlleAsnTYr 129  
 DB 298 CCGGCTCTCCGCGCATGGAGTGAACAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 357  
 OY 130 AAPPLeuLeuTYrAlIlleSeTyAlaLugLlYsProGlyGluLYsTYrGluLeuHlAlaAla 149  
 DB 358 GATCTCTTATATGATCATCTCCAACTGGGCGGCAAAAGTATGATGATGATGATGATGATG 417  
 OY 150 ThrAspProThProSeValValValHlSvalCYsGluSeAspGlnGluAsnGluGlu 169  
 DB 418 ACTGACACCACTCCAGCGT 474  
 OY 170 GluGluGluGluMetGluArgMetLysArgProLysProLYsIlleIlleGlnThrArgArg 189  
 DB 475 GAGAAAGAGAAATGAGAAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACGAGAGG 534

QY 190 ProGluTyrThrProIleHisLeuSer 198  
 |||||  
 Db 535 CCGGAGTACAGCGCGATCCACCTCAGC 561

RESULT 8  
 AAX01282  
 ID AAX01282 standard; cDNA to mRNA; 2174 BP.  
 XX  
 AC AAX01282;  
 XX  
 DT 09-APR-1999 (first entry)  
 XX  
 DE Human DSCR1 coding sequence.  
 XX  
 KW DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;  
 KW Central Nervous System development; mental retardation; heart defect; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 49..564  
 FT /tag= a  
 FT polyA\_signal 1541..1546  
 FT /tag= b  
 FT polyA\_signal 2132..2137  
 FT /tag= c  
 XX  
 PN US5869318-A.  
 XX  
 PD 09-FEB-1999.  
 XX  
 PP 07-JUN-1996; 96US-0665040.  
 XX  
 PR 07-JUN-1995; 95ES-0001140.  
 XX  
 PA (PALI/) PALLEJA X E.  
 XX  
 PI Fuentes JJ, Palleja XE, Pritchard M;  
 XX  
 DR WPI; 1999-152781/13.  
 DR P-PSDB; AAW73898.  
 XX

DNA encoding foetal brain proteins - believed to be associated with Down's syndrome

PS Claim 1; Column 15-18; 19pp; English.  
 XX

CC This sequence is the Down's Syndrome critical region 1 (DSCR1) gene of the invention. The DSCR1 gene was found to be located in the q22.1-22.2 region of human chromosome 21. An increase in the transient expression of DSCR1 mRNA in the brains of young rats, compared to expression levels in the brains of adult rats, suggests an important role for DSCR1 during the development of the Central Nervous System (CNS), and that over expression of DSCR1 may be involved in pathogenic abnormalities of mental retardation and/or heart defects as found in Down's syndrome patients.  
 XX

SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;  
 XX

Alignment Scores:  
 Pred. No.: 9,4e-86 Length: 2174  
 Score: 844.50 Matches: 161  
 Percent Similarity: 97.63% Conservative: 4  
 Best Local Similarity: 95.27% Mismatches: 3  
 Query Match: 80.97% Indels: 1  
 DB: 20 Gaps: 1

US-09-782-953-7 (1-198) x AAX01282 (1-2174)

QY 30 AlaLysPheGluSerLeuPheArgThrThrAspLysAspThrThrPheGlnTyrPheLys 49  
 |||||  
 Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTTCAGTATTITTAAG 117

QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
 |||||  
 Db 118 AGCTTCAAAACGAGTCAAGATAAACTTCAGCAACCCCTTCTCCGACGACATGCCAGGCTC 177

QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89  
 |||||  
 Db 178 CAGCTGCAATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTCTCTCAGACCTTA 237

QY 90 HisIleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerPro 109  
 |||||  
 Db 238 CACATAGGAAGCTCACACTGGCTCGCCAAATCCAGACCAAGCAGTTTCTGATCTCCCT 297

QY 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyr 129  
 |||||  
 Db 298 CCCGCTCTCCGCCAGTGGATGGAACAACATGGAAGATGCCACCCAGTCATAAATAT 357

QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149  
 |||||  
 Db 358 GATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGAAAGTATGAATTGCACGACGCG 417

QY 150 ThrAspProThrProSerValValHisValHisValCysGluSerAspGlnGluAsnGluGlu 169  
 |||||  
 Db 418 ACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTGCACAGAG---AAGGAG 474

QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189  
 |||||  
 Db 475 GAAGAAGAGGAAATGGAAGATGAGAGACCTAAGCCAAATAATTATCCAGACCAGGAGG 534

QY 190 ProGluTyrThrProIleHisLeuSer 198  
 |||||  
 Db 535 CCGGAGTACAGCGCGATCCACCTCAGC 561

RESULT 9  
 AAF18328  
 ID AAF18328 standard; DNA; 2358 BP.  
 XX  
 AC AAF18328;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polynucleotide sequence SEQ ID 347.  
 XX  
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055180-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PP 08-MAR-2000; 2000WO-US05918.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Ruben SM;  
 XX  
 DR WPI; 2000-587514/55.  
 DR P-PSDB; AAB58452.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 PS Claim 1; Page 805-806; 1425pp; English.  
 XX  
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytoskeletal; cardiovascular;  
 CC immunomodulatory; muscular active general; vulnery; gastrointestinal  
 CC general; nephrotropic; anti-infective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestional, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterization of the polynucleotide and protein  
 CC sequences.

XX Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

## Alignment Scores:

Pred. No.:	4,666-84	Length:	2358
Score:	830.00	Matches:	168
Percent Similarity:	88.89%	Conservative:	8
Best Local Similarity:	84.85%	Mismatches:	18
Query Match:	79.58%	Indels:	5
DA:	21	Gaps:	2

US-09-782-953-7 (1-198) x AAF18328 (1-2358)

OY 1 MetAapPhaArgAphSerTyraenPhaSerSerLeuLeuAlaCyValAlaAanAap 20  
 DB 123 GTGAGACCTCCAGACCTGCCCC-----AGGCCACCATGCGCTGACCTGACCGG 173  
 OY 21 AapValPhaSerGluSerGluThrArgAlaAlaPhaGluSerLeuPhaArgThrTyraap 40  
 DB 174 CCGGCTGTGTGACCGCCCTGCGGCCGCAATTGAGTCCCTCTTTAGAGATATGAC 233  
 OY 41 LyAapThrThaPhaGlnTyraPhaSerPhaLyAaArgValArgIleAanPhaSerAan 60  
 DB 234 AGGACATCACCTTTCAGATTTTAAAGCTTCAACGAGTCAAGATTAACCTCACGAC 293  
 OY 61 ProLeuSerAlaAlaAapAlaAaGluAaGluAaGluAaGluAaGluAaGluAaGlu 80  
 DB 294 CCTTCTCCGACGACATGCGACCTCCAGCTGCATTAAGCTGAGTTCTGGGAAAGAA 353  
 OY 81 MetLyAaLeuTyraPhaGlnThrLeuHisIleGlySerSerHisLeuAaLapProAan 100  
 DB 354 ATGAACTTATATTTCTCTCAGACCTTACACATGAAAGCTCACACTGGCTCCGCA-AAT 412  
 OY 101 ProAapLyGlnPhaLeuLeuLeuSerProProAlaSerProProValGlyTyraGlnVal 120  
 DB 413 CCAGACAAAGAGTTTGTGATCTCCCTCCGCTCCGSCAATGGGATGGAAACAAGTG 472  
 OY 121 GluAapAlaThrProValIleAaTyraPhaLeuLeuTyraAlaIleSerLeuGlyPro 140  
 DB 473 GAAGATGCGACCCAGCTCACTAACTATGATCTTATATGCCATCTCCAAAGCTGGGGCCA 532  
 OY 141 GlyGlnLyTyraGluLeuHisAlaAlaIleAaPProThaProSerValValIleHisVal 160  
 DB 533 GGGGAAAGATGATTAATGACGCGACGACTGACCACTCCAGCTGTGTGTCCATGTA 592  
 OY 161 CyGluSerAapGlnGluAanGluGluGluGluGluGluGluGluGluGluGluGlu 180  
 DB 593 TGTAAAGATGATCAAGG---AAGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 649  
 OY 181 LyAaPProLyIleIleGlnThrArgAaGProGluTyraThrIleHisLeuSer 198  
 DB 650 AAGCCAAAATATATCCAGACGAGGCGCGAGTACACGCCGATCCACCTCAACG 703  
 RESULT 10  
 AAD30154  
 ID AAD30154 standard; DNA; 594 BP.

XX AC AAD30154;  
 XX XX 17-MAY-2002 (first entry)  
 DT 17-MAY-2002 (first entry)  
 XX XX Mouse MCIP associated DNA #4.  
 DE Mouse MCIP associated DNA #4.  
 XX XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; mouse; gene; ds.  
 XX XX Mus musculus.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT CDS 1..594  
 FT /tag= a  
 FT /product= "Mouse MCIP associated protein #4"  
 XX XX MO200204491-A2.  
 XX XX 17-JAN-2002.  
 PD 17-JAN-2002.  
 XX XX 06-JUL-2001; 2001WO-US21662.  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX XX 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 ER 13-FEB-2001; 2001US-0782953.  
 XX XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX XX Williams SR, Rothermel B;  
 PI WI: 2002-179698/23.  
 DR P-PSDB; AAE18913.  
 XX XX  
 DR Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 PT  
 PS Disclosure; Page 153-154; 174pp; English.  
 XX XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated DNA.  
 CC Note: This sequence has been described as mouse MCIP1 in the  
 CC specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.  
 CC  
 XX  
 SQ Sequence 594 BP; 148 A; 165 C; 149 G; 132 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,886-62 Length: 594  
 Score: 632.50 Matches: 121  
 Percent Similarity: 77.13% Conserv: 24  
 Best Local Similarity: 64.36% Mismatches: 34  
 Query Match: 60.64% Indels: 9  
 DB: 24 Gaps: 2  
 US-09-782-953-7 (1-198) x AAD30154 (1-594)  
 OY 11 SerSerLeuIleAlaCyValAlaAanAapValPhaSerGluSerGluThrArgAla 30  
 DB 31 TCCACTCTGTGCTGT 90



Db 614 GTGACGCTGTGACATGACATAGAGAGAGAGACCAAG----- 658  
 Oy 178 LysArgProLysProLysIleIleGlnThrArgArgProGluTyrThrPro 194  
 Db 659 ACTTCCCAAGCCAAATATCATCTCAAACTGGCGCTCGCTGCCACCC 709

RESULT 12  
 ABA91457  
 ID ABA91457 standard; cDNA; 3159 BP.  
 AC ABA91457;  
 DT 18-APR-2002 (first entry)  
 DE Human Down syndrome critical region 1-like 1 protein cDNA.  
 KM Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;  
 KM Down syndrome; Alzheimer's disease; dementia; transgenic;  
 KM neuroprotective; noctropic; anticonvulsant; diagnosis;  
 KM gene therapy; gene; 88.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 185..952  
 FT /\*tag= a  
 FT /product= "DSCR1L1"  
 FT /transl\_except= (pos:215..217, aa:Xaa)  
 FT /note= "Xaa = unknown"  
 PT  
 PN WO200204513-A2.  
 PD 17-JAN-2002.  
 PP 11-JUL-2001; 2001WO-US21982.  
 PR 11-JUL-2000; 2000US-0614474.  
 PA (INCY-) INCYTE GENOMICS INC.  
 PI Loring JF, Tingley DM, Edwards CM, Streeter DG;  
 DR WPI; 2002-164633/21.  
 DR P-PSDB; AAM50760.  
 PT Novel Down syndrome critical region 1-like 1 protein and nucleic acid  
 PT encoding the protein useful for diagnosis and treatment of Alzheimer's  
 PT disease, Down syndrome and other forms of dementia -  
 PS Claim 2; Page 46-47; 54pp; English.

The present sequence is that of cDNA clone incyte ID No: 247500.5  
 encoding novel Down syndrome critical region 1-like 1 protein  
 (DSCR1L1 alpha, see AAM50760). Northern analysis indicated  
 expression of DSCR1L1 alpha in various libraries, with the highest  
 abundance in tissues from the nervous system, including tissues  
 associated with schizophrenia, Huntington's disease, epilepsy and  
 amyotrophic lateral sclerosis. An absence of DSCR1L1 expression  
 was observed in 7 of 8 libraries from subjects with Alzheimer's  
 disease. A claimed method for detecting differential expression of  
 a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down  
 CC syndrome, Alzheimer's disease and other forms of dementia. A probe  
 CC from such a nucleic acid is useful for identifying naturally  
 CC occurring molecules encoding DSCR1L1 alpha, allelic variants or  
 CC other molecules. The nucleic acids are useful for producing  
 CC transgenic cell lines or organisms which model human disorders.  
 CC They may also be used in gene therapy, and to screen for ligands  
 CC which specifically bind the nucleic acid molecule, such as  
 CC a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,  
 CC peptide, transcription factor, repressor or regulatory molecule,  
 CC for use as a therapeutic.  
 XX  
 XX Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;

Alignment Scores:  
 Pred. No.: 2,266-60  
 Score: 623.00  
 Percent Similarity: 73.10%  
 Best Local Similarity: 62.44%  
 Query Match: 59.73%  
 DB: 24  
 Gaps: 3

US-09-782-953-7 (1-198) x ABA91457 (1-3159)

Oy 2 AspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAspAsp 21  
 Db 371 GACTTCATGACCTCCCC-----AACTCGTTGTTGGTGCATGTTCCACCTCA 421  
 Oy 22 ValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgTyrTyrAspLys 41  
 Db 422 GTGTTTGAAG 481  
 Oy 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61  
 Db 482 TGGTGACGTTCCAGCTATTATTAAGAGTTTCAGACGTGCTCGTATTAACCTTCAGCAATCCT 541  
 Oy 62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81  
 Db 542 AAATCGACGCCCGAGCTAGATGAGCTTCATGAAACCAATTCAGAGGAGAGAGAGAGAGATTA 601  
 Oy 82 LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSer-----HisLeuAla 97  
 Db 602 AAGCTCTACTTTGGACAGAGCTTCAGACTCCAGACAGATGAGAGAGAGAGAGAGAGAGAGAGAG 661  
 Oy 98 ProProAsnProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyr 117  
 Db 662 CCACCCGACCTGCGCAACAGTTTCTATCTGCCCCCTCTCTCCCACTGTTGGCTGG 721  
 Oy 118 LysGlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLys 137  
 Db 722 CAGCCCATCAACGATGCCAGCCGCTCTCAACTATGACTCTCTATGCTGTGGCCAA 781  
 Oy 138 LeuGlyProGlyGlyLysTyrGlyLeuHisAlaIleThrAspProThrProSerValVal 157  
 Db 782 CTAGGACCCAGGAGAGAGATATGAGCTCCATGACAGGAGCTGACCCCAAGTGTCTGT 841  
 Oy 158 ValHisValCysGluSerAspGlnGluAsnGluGluGluGluGluGluMet 177  
 Db 842 GTGACGCTGTGACATGACATGAG 886  
 Oy 178 LysArgProLysProLysIleIleGlnThrArgArgProGluTyrThrPro 194  
 Db 887 ACTTCCCAAGCCAAATATCATCTCAAACTGGCGCTCGCTGCCACCC 937

RESULT 13  
 AAD30157  
 ID AAD30157 standard; DNA; 3184 BP.  
 XX  
 AC AAD30157;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated DNA #3.  
 DE  
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KM heart failure; cardiomyopathy; heart disease; human; gene; de.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 205..783  
 FT /\*tag= a  
 FT /product= "Human MCIP associated protein #3"  
 PT  
 PN WO200204491-A2.

PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 XX MPI: 2002-179698/23.  
 DR P-PSDB; AAE18916.  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding; expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 163-165; 174pp; English.  
 XX  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence has been described for determining the  
 CC Note: This sequence is human MCIP associated DNA.  
 CC specification, however the sequence seems to be a polynucleotide encoding  
 CC a MCIP associated protein.  
 XX  
 SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.63e-59 Length: 3184  
 Score: 615.50 Matches: 119  
 Percent Similarity: 75.53% Conservative: 23  
 Best Local Similarity: 63.30% Mismatches: 37  
 Query Match: 59.01% Indels: 9  
 DB: 24 Gaps: 2  
 US-09-782-953-7 (1-198) x AAD30157 (1-3184)  
 QY 11 SerSerLeuIleAlaCysValAlaAsnAspValPheSerGluSerGluThrArgAla 30  
 DB 220 TCCACTCTGGTTCCTGCTGGTGGATGTCAGGCTCTTTACCAATCAGGAGGTTAAGGAA 279  
 QY 31 LysPheGluSerLeuPheArgThrThrAspLysAspThrPheGlnTyrPheLysSer 50  
 DB 280 AAATTTGGGGGAGCTGTTTCGGACTTATGATGACTGTGTGACGCTTCAGCTATTTAAGACT 339  
 QY 51 PheLysArgValArgIleAlaPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 DB 340 TTCAGAGGTGTCGTATTAACCTTACCAATCTTAATCTGACGCCGAGCTAGGATAGAG 399  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 DB 400 CTTTCATGAAACCAATTCAGAGGGAATAAATAAGCTCTACTTTGCACAGGTTTCAGACT 459  
 QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106  
 DB 460 CCAGACAGACATGGAGACAAATCGCACTTGGCTCCACCCAGCCTGCCAAACAGTTTCTC 519  
 QY 107 IleSerProProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProVal 126  
 DB 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.

DB 520 ATCTGCCCCCTTCCTCCCCACCTGTTAGTGGCAGCCCATCAACGATGCCACGCCAGTC 579  
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGlnLysTyrGluLeu 146  
 DB 580 CTCACACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGAGAGATATGAGCTC 639  
 QY 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166  
 DB 640 CATGCAGGAGCTGAGTCCACCCCAAGTGTGCTGTCGACGTCGTCGACAGTGCATAGAG 699  
 QY 167 AsnGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGln 186  
 DB 700 GAAGAAGAGGACCAAG-----ACTTCCCAAGCAAAATCATCCAA 744  
 QY 187 ThrArgArgProGluTyrThrPro 194  
 DB 745 ACTCGGCTCTGGCTGCCACCC 768  
 RESULT 14  
 ABL61768  
 ID ABL61768 standard; DNA; 3184 BP.  
 XX  
 AC ABL61768;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:105.  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 26-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.





PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 PA (AVAL-) AVALON PHARM.  
 XX  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,  
 PI Soppet DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 XX  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 XX  
 PS Claim 1; SEQ ID 3524; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 XX  
 SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.63e-59 Length: 3184  
 Score: 615.50 Matches: 119  
 Percent Similarity: 75.53% Conservative: 23  
 Best Local Similarity: 63.30% Mismatches: 37  
 Query Match: 59.01% Indels: 9  
 DB: 24 Gaps: 2

US-09-782-953-7 (1-198) x ABL65187 (1-3184)

QY 11 SerSerLeuLeuAlaCysValAlaAsnAspValPheSerGluThrArgAla 30  
 Db 220 TCACCTCTGGTTCCTGCTGTGGTGGATGTCGAGGTCTTTACCAATCAGGAGGTTAAGGAA 279  
 QY 31 LysPheGluSerLeuPheArgThrTyraPheLysAspThrPheGlnTyrPheLysSer 50  
 Db 280 AATTTGGGGGAGCTGTTCCGACTTATGACTGTGTGACGTTCCAGCTATTTAAGNGT 339  
 QY 51 PheLysArgValArgLeuAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 Db 340 TTCAGACGTGTCGTATAAATTCAGCAATCCTAATCTGCAGCCCGAGCTAGGATAGAG 399  
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 Db 400 CTTTCATGAACCAATTCAGAGGGGAAAAAATTAAAGCTCTACTTTGCACAGGTTTCAGACT 459

QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106  
 Db 460 CCAGAGACAGATGCAGACAAACTGGCTCCACCCAGCTGCCAAACAGTTTCTC 519  
 QY 107 IleSerProProAlaSerProProValGlyTTPlyGlnValGluAspAlaThrProVal 126  
 Db 520 ATCTCGCCCTTCTCTCCACCTGTTAGCTGCAGCCCATCAACGATGCCAGCCAGTC 579  
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146  
 Db 580 CTCAACTATGACCTCTCTATGCTGTGGCCAAACTAGGACCCAGAGAGATATGAGCTC 639  
 QY 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166  
 Db 640 CATGCAGGAGACTGAGTCCACCCCAAGTGTCTGTCACGTGTGCAGACAGTGCATAGAG 699  
 QY 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysLysLysLys 186  
 Db 700 GAAGAAGAGGAGCCCAAG-----ACTTCCCAAGCCAAATCATCCAA 744  
 QY 187 ThrArgArgProGluTyrThrPro 194  
 Db 745 ACTCGGCTCTGCGCTGCACCC 768

Search completed: December 14, 2002, 20:09:05  
 Job time : 145.568 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 1079.63 Seconds  
(without alignments)  
16146.861 Million cell updates/sec

Title: US-09-782-953-1  
Perfect score: 599  
Sequence: 1 gagggcgaagaagactcca.....gaggaggaagaggaagagat 599

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.on.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	599	100.0	599	6	AX365311	AX365311 Sequence
2	573	95.7	2198	10	BC013551	BC013551 Mus muscu
3	524.4	87.5	2141	10	AF263239	AF263239 Mus muscu
4	504.4	84.2	597	6	AX365315	AX365315 Sequence
5	504.4	84.2	597	10	AF237790	AF237790 Mus muscu
6	453.2	75.7	2216	10	CGU60263	U60263 Cricetus
7	435.8	72.8	626	10	AB075973	AB075973 Rattus no
8	427.8	71.4	621	10	AF282255	AF282255 Mus muscu
9	427.8	71.4	2125	10	AF260717	AF260717 Mus muscu
10	426.2	71.2	597	6	AX365312	AX365312 Sequence
11	426.2	71.2	597	10	AF237789	AF237789 Mus muscu
12	425.4	71.0	2224	10	AF263240	AF263240 Mus muscu
13	413.8	69.1	2331	6	AX365321	AX365321 Sequence
14	413.8	69.1	2346	9	HSU85267	HSU85267 Homo sapien
15	404.8	67.6	2348	6	AX281651	AX281651 Sequence
16	394	65.8	562	9	HSU53821	U53821 Homo sapien
17	349.2	58.3	2212	6	AX365324	AX365324 Sequence
18	349.2	58.3	2227	9	HSU85266	U85266 Homo sapien
19	349.2	58.3	2289	9	BC002864	BC002864 Homo sapi
20	349.2	58.3	2407	9	AK092184	AK092184 Homo sapi
21	343.8	57.4	2284	9	HSU85265	U85265 Homo sapien
22	341.8	57.1	2173	6	AX410694	AX410694 Sequence
23	341.8	57.1	2173	9	HSU28833	U28833 Homo sapien
24	341.2	57.0	2679	9	AK055845	AK055845 Homo sapi
25	340.2	56.8	798	9	AF400429	AF400429 Homo sapi
26	340.2	56.8	2174	6	AR034241	AR034241 Sequence
27	202.2	33.8	3184	6	AX329596	AX329596 Sequence
28	202.2	33.8	3184	6	AX333015	AX333015 Sequence
29	202.2	33.8	3184	6	AX365327	AX365327 Sequence
30	202.2	33.8	3184	6	AX420436	AX420436 Sequence
31	202.2	33.8	3184	9	HUM2AK14	D83407 ZAK1-4 mRNA
32	197.6	33.0	934	6	AX074351	AX074351 Sequence
33	197.6	33.0	3159	6	AX420425	AX420425 Sequence
34	197.6	33.0	3261	9	AK090990	AK090990 Homo sapi
35	194.4	32.5	3240	9	AY034085	AY034085 Homo sapi
36	194.4	32.5	3253	9	AY034086	AY034086 Homo sapi
37	191	31.9	594	6	AX365318	AX365318 Sequence
38	191	31.9	594	10	AF237791	AF237791 Mus muscu
39	191	31.9	659	10	AF237887	AF237887 Mus muscu
40	189.8	31.7	720	6	AX365333	AX365333 Sequence
41	189.8	31.7	720	9	AF176117	AF176117 Homo sapi
42	189.4	31.6	3227	10	AB061524	AB061524 Mus muscu
43	188.4	31.5	828	6	AX365330	AX365330 Sequence
44	188.4	31.5	828	6	AX420437	AX420437 Sequence
45	188.4	31.5	828	9	AF176116	AF176116 Homo sapi

ALIGNMENTS

RESULT 1  
AX365311  
LOCUS AX365311 599 bp DNA linear PAT 15-FEB-2002  
DEFINITION Sequence 1 from Patent WO2004491.  
ACCESSION AX365311  
VERSION AX365311.1 GI:18697042  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Williams, S.R. and Rothermel, B.  
TITLE Methods and compositions relating to muscle selective calcineurin  
interacting protein (mcip)



```

Db 421 TGGCTGGAAACAGTAGAGATGCCACCCCGTCATAAATTACGATCTTTTATATGCCAT 480
Qy 483 CTCAAGCTGGGCGCAGAGAGAGTATGAAGTCTGATGCGAGCGACACACCACTCCAG 542
Db 481 CTCAAGCTGGGCGCAGAGAGAGTATGAAGTCTGATGCGAGCGACACCACTCCAG 540
Qy 543 TGTGCTGCTCAGCTGTGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAT 599
Db 541 TGTGCTGCTCAGCTGTGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAT 597

RESULT 3
AF263239 2141 bp mRNA linear ROD 30-MAY-2000
LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
DEFINITION alternatively spliced.
ACCESSION AF263239
VERSION AF263239.1 GI:8102011
KEYWORDS
SOURCE
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLES Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.
JOURNAL Down syndrome candidate region 1 (Dscr1), one of three
REFERENCE alternatively spliced exon 1 transcripts
AUTHORS 2 (bases 1 to 2141)
TITLES Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.
JOURNAL Direct Submission
TITLES Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
JOURNAL Avia. Castelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona
08907, Spain
FEATURES
source Location/Qualifiers
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/note="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"
/translation="MHFRDYSYNFSSLIACVANDVFESETRAKFESLFRDYDKDTT
FOYKFSKRVIRNFSPNSAADARLRLHKTFLGKMKLYFAOTLHIGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLLYAISKLGPGEKYLHAATDTPSVVVHV
CESDQNEEBEEMERMKPKKIOTRRPEYTPHLS"
BASE COUNT 505 a 527 c 558 g 548 t 3 others
ORIGIN
Query Match 87.5%; Score 524.4; DB 10; Length 2141;
Best Local Similarity 99.5%; Pred. No. 1.2e-155;
Matches 547; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy 52 GTTAAGCGTCTGCCCGTGAAGAGAGATG-ATTTTAGGCACTTTAGTACAATTTTA 110
Db 2 GTTAAGCGTCTGCCCGTGAAGAGAGATGCAATTTTAGGCACTATAGTACAATTTTA 61
Qy 111 GTCCTCGTATGCTGTGTGGCAACAGATGCTCTTCAGCGAAAGTGAGACCGGCCA 170
Db 62 GTCCTCGTATGCTGTGTGGCAACAGATGATGCTCTTCAGCGAAAGTGAGACCGGCCA 121
Qy 171 AATTTGAATCCCTCTTCAGAAATATGACAGACACCACTTCCAGTATTTTAGAGCT 230
Db 122 AATTTGAATCCCTCTTCAGAAATATGACAGACACCACTTCCAGTATTTTAGAGCT 181
Qy 231 TCAAGCTGTCGGGATAAATTTCAGCAACCCCTTATGTCAGCCGATGCCAGGCTGGCG 290
Db 182 TCAAGCTGTCGGGATAAATTTCAGCAACCCCTTATGTCAGCCGATGCCAGGCTGGCG 241

```

```

Qy 291 TGCACAAGACCGAGTTCTCTGGGGAAGGAAATCAAGTTGTTATTTTCTCAGACTTTACACA 350
Db 242 TGCACAAGACCGAGTTCTCTGGGGAAGGAAATCAAGTTGTTATTTTCTCAGACTTTACACA 301
Qy 351 TAGGAAGTTTACACCTTGGCTCCG-CCAAATCCCGCAAAACAGTTCTCTCATCTCCCTCCGG 409
Db 302 TAGGAAGTTTACACCTTGGCTCCGCCCCAATCCCGCAAAACAGTTCTCTCATCTCCCTCCGG 361
Qy 410 CCTCTCTCCCTCCCTGGCTGGTGAACAAGTAGAAGATGCCACCCCGTCTATAAATTAGCATC 469
Db 362 CCTCTCTCCCTCCCTGGCTGGTGAACAAGTAGAAGATGCCACCCCGTCTATAAATTAGCATC 421
Qy 470 TTTTATATGCCATCTCCCAAGCTGGGCGCAGGAGAGAGTATGAAGTTCGATGCGAGCGACAG 529
Db 422 TTTTATATGCCATCTCCCAAGCTGGGCGCAGGAGAGAGTATGAAGTTCGATGCGAGCGACAG 481
Qy 530 ACACCACTCCCAAGTGTGTGCTCCACGTGTGTGAGAGTGACCAAGAGATGAGGAGGAAG 589
Db 482 ACACCACTCCCAAGTGTGTGCTCCACGTGTGTGAGAGTGACCAAGAGATGAGGAGGAAG 541
Qy 590 AGGAAGAGAT 599
Db 542 AGGAAGAGAT 551

RESULT 4
AX365315 597 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 5 from Patent WO0204491.
DEFINITION AX365315
ACCESSION AX365315
VERSION AX365315.1 GI:18697045
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLES Williams,S.R. and Rothermel,B.
JOURNAL Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
PATENT Patent: WO 0204491-A 5 17-JAN-2002;
BOARD Board of Regents, The University of Texas System (US); Williams,
SANDERS Sanders R. (US); Rothermel, Beverly (US)
FEATURES
source Location/Qualifiers
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
1..597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23809.1"
/db_xref="GI:18697046"
/translation="MDPRDYSYNFSSLIACVANDVFESETRAKFESLFRDYDKDTT
FOYKFSKRVIRNFSPNSAADARLRLHKTFLGKMKLYFAOTLHIGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLLYAISKLGPGEKYLHAATDTPSVVVHV
CESDQNEEBEEMERMKPKKIOTRRPEYTPHLS"
BASE COUNT 170 a 156 c 142 g 129 t
ORIGIN
Query Match 84.2%; Score 504.4; DB 6; Length 597;
Best Local Similarity 99.6%; Pred. No. 2.2e-149;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 83 GATTTAGGCACTTTAGTACAATTTTAGCTCCCTGATTTGCTTGTGGCAACAGATGAT 142
Db 4 GATTTAGGCACTTTAGTACAATTTTAGCTCCCTGATTTGCTTGTGGCAACAGATGAT 63
Qy 143 GTCCTTCAGCGAAAGTGAGACCGAGCGCAAAATTTGAATCCCTCTTCAGAAATATGACAAG 202
Db 64 GTCCTTCAGCGAAAGTGAGACCGAGCGCAAAATTTGAATCCCTCTTCAGAAATATGACAAG 123
Qy 203 GACACCACTTCCAGTATTTTAAAGAGCTTCAAAACGTGTCGGGATAAACTTCAGCAACCCC 262

```

```

|||||
Db 124 GACACCACTTCCAGTATTTTAAAGCTTCAAAACGTCCGATAACTTCAGCAACCCC 183
Oy 263 TTATCTGACCCGATGCCAGGCTGCGCTGCACAACAAGCTTCCGTGGGAAGAAATG 322
Db 184 TTATCTGACCCGATGCCAGGCTGCGCTGCACAACAAGCTTCCGTGGGAAGAAATG 243
Oy 323 AAGTTGTATTTTGTCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCG-CCATGCC 381
Db 244 AAGTTGTATTTTGTCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCG-CCATGCC 303
Oy 382 GACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGGCTGGAAGAAAGTAGAA 441
Db 304 GACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGGCTGGAAGAAAGTAGAA 363
Oy 442 GATGCCACCCCGCTCATTAATAGATCTTTATATGCAATCTCCAAAGTGGGGCCAGGA 501
Db 364 GATGCCACCCCGCTCATTAATAGATCTTTATATGCAATCTCCAAAGTGGGGCCAGGA 423
Oy 502 GAGAAGTATGAACTGCATGCAGCAGACAGACACACACTCCAGTGTGGTCCACGTTGT 561
Db 424 GAGAAGTATGAACTGCATGCAGCAGACAGACACACACTCCAGTGTGGTCCACGTTGT 483
Oy 562 GAGAGTGACCAAGAGATGAGAGAGAGAGAGAT 599
Db 484 GAGAGTGACCAAGAGATGAGAGAGAGAGAGAT 521

RESULT 5
AF237790 597 bp mRNA linear ROD 12-APR-2000
LOCUS Mus musculus myocyte-enriched calcineurin interactor protein 1
DEFINITION AF237790
AF237790.1 GI:7542528
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 597)
Rothermel,B.A., Vega,R.B., Yang,J., Wu,H., Baszel-Duby,R.S. and
Williams,R.S.
2 (bases 1 to 597)
J. Biol. Chem. (2000) in press
3 (bases 1 to 597)
Rothermel,B.A., Vega,R.B., Yang,J., Wu,H., Baszel-Duby,R.S. and
Williams,R.S.
Direct Submission
Submitted (23-PEB-2000) Internal Medicine, University of Texas
Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX
75390, USA
FEATURES
source
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..597
/notes="similar to the EST with Genbank Accession Number
AA200984; similar to Homo sapiens DSCR1"
AA200984; similar to Homo sapiens DSCR1"
/codon_start=1
/product="myocyte-enriched calcineurin interactor protein
1 splice variant 4"
/protein_id="AA63486.1"
/db_xref="GI:7542528"
/translacion="MDFRDPSTYFSSLIACVANDVFSSESTRAKESLFRYDKDTT
FOYKSPKRVINSPNLSADARLRJHTEFLGKMKLYFAQTLLIGSSHLAPPPD
KORLISPPAPVGMKOVEDATPVINYDLTALYSKGPGEKELHATPTPSVVVHV
CESDQNEEESEEMERKRPKRIQTRREYPIHLS"
CDS
170 a 156 c 142 g 129 c
BASE COUNT
ORIGIN

```

```

Query Match 84.2%; Score 504.4; DB 10; Length 597;
Best Local Similarity 99.6%; Pred. No. 2,2e-149;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 83 GATTTAGGAGCTTAGCTACATATTTTAACTCCCTGATGCTGTGTGGCAAGAGATGAT 142
Db 4 GATTTAGGAGCTTAGCTACATATTTTAACTCCCTGATGCTGTGTGGCAAGAGATGAT 63
Oy 143 GTCTTCAGCCGAAAGTGAAGACAGGAGCCAAATTTGAATCCCTTCAGAACATATGACAG 202
Db 64 GTCTTCAGCCGAAAGTGAAGACAGGAGCCAAATTTGAATCCCTTCAGAACATATGACAG 123
Oy 203 GACACCACTTCCAGTATTTTAAAGCTTCAACAGTGTCCGATAACTTCAGAACCCC 262
Db 124 GACACCACTTCCAGTATTTTAAAGCTTCAACAGTGTCCGATAACTTCAGAACCCC 183
Oy 263 TTATCTGACCCGATGCCAGGCTGCGCTGCACAACAAGCTTCCGTGGGAAGAAATG 322
Db 184 TTATCTGACCCGATGCCAGGCTGCGCTGCACAACAAGCTTCCGTGGGAAGAAATG 243
Oy 323 AAGTTGTATTTTGTCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCG-CCATGCC 381
Db 244 AAGTTGTATTTTGTCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCG-CCATGCC 303
Oy 382 GACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGGCTGGAAGAAAGTAGAA 441
Db 304 GACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTGGCTGGAAGAAAGTAGAA 363
Oy 442 GATGCCACCCCGCTCATTAATAGATCTTTATATGCAATCTCCAAAGTGGGGCCAGGA 501
Db 364 GATGCCACCCCGCTCATTAATAGATCTTTATATGCAATCTCCAAAGTGGGGCCAGGA 423
Oy 502 GAGAAGTATGAACTGCATGCAGCAGACAGACACACACTCCAGTGTGGTCCACGTTGT 561
Db 424 GAGAAGTATGAACTGCATGCAGCAGACAGACACACACTCCAGTGTGGTCCACGTTGT 483
Oy 562 GAGAGTGACCAAGAGATGAGAGAGAGAGAGAT 599
Db 484 GAGAGTGACCAAGAGATGAGAGAGAGAGAGAT 521

RESULT 6
CGU60263 2216 bp mRNA linear ROD 03-SEP-1997
LOCUS Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.
DEFINITION CGU60263
CGU60263
U60263.1 GI:2351390
VERSION
KEYWORDS
SOURCE
ORGANISM
Cricetulus griseus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 2216)
Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and
Davies,K.J.
Hamster adapt78 mRNA is a Down syndrome critical region homologue
that is inducible by oxidative stress
Arch. Biochem. Biophys. 342 (1), 6-12 (1997)
97329055
MEDLINE
JOURNAL
PUBMED
2 (bases 1 to 2216)
Crawford,D.R., Leahy,K.L. and Davies,K.J.A.
Direct Submission
Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical
College, 47 New Scotland Avenue, Albany, NY 12208, USA
FEATURES
source
1..2216
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/cell_type="HA-1"
1..2216
/gene="adapt78"

```

CDS	70..663	/gene="adapt78"	
		/note="similar to human DSCR1; an oxidant- and calcium-inducible mRNA"	
		/codon_start=1	
		/product="Adapt78"	
		/protein_id="AA068517.1"	
		/db_xref="GI:23513191"	
		/translation="MHPRDFNYNFSLLIACVANGDVSESETRAKFESLFRVYDKDIT FQYKSPKRVIRINFSNLSAADARLQLHKTEFLGKEMKLYFAOTLHGSSHLAPPND KQFLISPPAPVPVGMKQVEDATPVINDLYALSKLPGKEKYLHAATDTTSPVVHVH CESDQNEEBEEMERMKRPKIKIOTRRPEYTPHLS"	
BASE COUNT	544 a 515 c 564 g 593 t		
ORIGIN			
Query Match	75.7%; Score 453.2; DB 10; Length 2216;		
Best Local Similarity	90.2%; Pred. No. 6.4e-133;		
Matches	507; Conservative 0; Mismatches 53; Indels 2; Gaps 2;		
Qy	34	TGACAGAGCGAGTCGTTCCGTTAGCGCTGCCCCGCTGAAGAACGAGATG-ATTTAGGG	92
Db	23	TGACAGAGAGTCGTGAGTCATTTAGCGTCTGCCCTGTGAAGAACGAGATGCAATTTAGGG	82
Qy	93	ACTTTAGCTACAAATTTTAGCTCCCTGATTCCTGTTGGCAACGATGATCTTCAGCG	152
Db	83	ACTTTAACTACAATTTTAGCTCCCTGATTCGCTGTGGCAACGCTGATGCTTCAGCG	142
Qy	153	AAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAGCACCACT	212
Db	143	AAAGTGAAACCGAGCGCCAAATTTGAATCCCTCTTCAGAGCTATGACAGGACATCACT	202
Qy	213	TCCAGTATTTTAAAGAGCTTCAAAACGCTCCGGATTAACCTTCAGCAACCCCTTATCTGCAG	272
Db	203	TCCAGTATTTTAAAGAGCTTCAAAACGCTTCCGAATAAATTCAGCAACCCCTTATCCGCAG	262
Qy	273	CCGATGCCAGGCTCGCGCTGCACAAGACCGAGTTCTCTGGGGAAGGAAATGAAGTCTGATT	332
Db	263	CTGACGCCAGGCTGCAGCTGCATAGACGAGTTCCTCTGGGGAAGGAAATGAAGCTGACT	322
Qy	333	TTGCTCAGAGCTTTACACATAGGAAGTTCACCTGGCTCCG-CCAATCCGACAAACAGT	391
Db	323	TTGCTCAGAGCTTTACACATAGGAAGTTCACCTGGCTCCGCGCCCAATCCAGACAGCACT	382
Qy	392	TCCTCATCTCCCTCCGCGCTCTCTCCGCTTGGCTGGGAAACAAGTAGAGATGCCACCC	451
Db	383	TCCTCATCTCTCTCCGCGCTCTCCACAGTTGGTTGGGAAGCAAGTAGAAGAGCTACCC	442
Qy	452	CCGTCATAAATTACGATCTTTATATGCAATCTCCAGCTGGGGCCAGGAGAGAAAGTATG	511
Db	443	CAGTCATAAATTACGATCTTTATATGCTATCTCCAAAGCTGGGGCCAGGTGAGAGTATG	502
Qy	512	AACGTCATGCAGCGACAGACACCTCCAGTGTGGTGTGCTCCAGCTGTGTGAGATGACC	571
Db	503	AGCTGATCGACGACAGACACCTCCAGTGTGGTGTGCTCCAGCTGTGTGAGATGACC	562
Qy	572	AAGAGAAATGAGGAGGAGGAGGA 593	
Db	563	AAGAGAAATGAGGAGGAGGAGGA 584	
RESULT 7			
AB075973			
LOCUS	AB075973	626 bp mRNA linear	ROD 27-JUL-2002
DEFINITION	Rattus norvegicus mcip 1 mRNA for myocyte-enriched calcineurin-interacting protein 1, complete cds.		
ACCESSION	AB075973		
VERSION	AB075973.1	GI:21998843	
KEYWORDS	Rattus norvegicus cDNA to mRNA.		
SOURCE	Rattus norvegicus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1		
AUTHORS	Odashima,M., Nagata,K., Obata,K., Somura,F., Izawa,H., Miyazaki,T., Murata,Y., Seo,H., Yamada,Y. and Yokota,M.		
TITLE	rat myocyte-enriched calcineurin interactin protein 1, splice variant 4 mRNA, complete cds		
JOURNAL	Unpublished		
AUTHORS	2 (bases 1 to 626)		
TITLE	Obata,K., Nagata,K., Odashima,M., Somura,F., Yamada,Y. and Yokota,M.		
JOURNAL	Submitted (11-DEC-2001) Koji Obata, Nagoya University, Department of Clinical Pathophysiology; 65 Teurumai-cho, Showa-ku, Nagoya, Aichi 466-8550. Japan (E-mail:obata@etsu.med.nagoya-u.ac.jp, Tel:81-52-744-2577, Fax:81-52-744-2977)		
FEATURES	Location/Qualifiers		
source	1..626		
gene	/organism="Rattus norvegicus"		
CDS	1..626		
	/gene="mcip 1"		
	19..618		
	/gene="mcip 1"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin-interacting protein 1"		
	/protein_id="BAC06443.1"		
	/db_xref="GI:21998844"		
	/translation="MHPRDFNYNFSLLIACVANGDVSESETRAKFESLFRVYDKDIT FQYKSPKRVIRINFSNLSAADARLQLHKTEFLGKEMKLYFAOTLHGSSHLAPPND KQFLISPPAPVPVGMKQVEDATPVINDLYALSKLPGKEKYLHAATDTTSPVVHVH CESDQNEEBEEMERMKRPKIKIOTRRPEYTPHLS"		
BASE COUNT	182 a 163 c 154 g 127 t		
ORIGIN			
Query Match	72.8%; Score 435.8; DB 10; Length 626;		
Best Local Similarity	91.7%; Pred. No. 1.7e-127;		
Matches	472; Conservative 0; Mismatches 42; Indels 1; Gaps 1;		
Qy	84	ATTTTAGGAGCTTTAGCTACAAATTTTAGCTCCCTGATGCTTGTGTGGCAACGATGATG	143
Db	23	ATTTTAGGAGCTTTAACTACAAATTTTAGCTCCCTGATGCTTGTGTGGCAACGATGATG	82
Qy	144	TCCTCAGCAAGAGTGAGACCGAGGCCAAATTTGAATCCCTCTTCAGAACATATGACAGG	203
Db	83	TCCTCAGCAAGAGTGAGACCGAGGCCAAATTTGAATCCCTCTTCAGACGATGACAGG	142
Qy	204	ACACACCTTCCAGTATTTTAAGAGCTTCAAAAGTGTCCGATAACTTCAGCAACCCCT	263
Db	143	ACATCACCTTCCAGTATTTTAAGAGCTTCAAGCGGTGTCGAATAAACTTCAGCAACCCCT	202
Qy	264	TATCTGAGCGCGATGCCAGCTGCGCTGCACAGACCGAGTTCCTGGGGAAGGAAATGA	323
Db	203	TGCTGGGCGGACCGCAGCTGCGCTGCATAGACAGAGTTCCTGGGGAAGGAGATGA	262
Qy	324	AGTTGTATTTGTCTAGACTTTACACATAGGAAGTTACACCTGGCTCCG-CCAATCCCG	382
Db	263	AACGTACTTTGCCAGACTTTACACATAGGAAGTTACACCTGGCTCCGCCCAATCCAG	322
Qy	383	ACAAACAGTTTCTCATCTCCCTCCGCTCTCTCCGCTCTCTCCGTTGGCTGGAACAGTGAAG	442
Db	323	ACAAACAGTTTCTCATCTCTCCCTCCGCTCTCTCCGCTCTCTCCGCTGGAACAGTGAAG	382
Qy	443	ATGCCACCCCGCTCATAAATTTAGCATCTTTATATGCCATCTCCAGCTGGGCCAGGAG	502
Db	383	ACGCTACCCCGCTCATAAATTTAGCATCTTTATATGCCATCTCCAGCTGGGCCAGGAG	442
Qy	503	AGAAGTATGAATGCAATGACGAGACAGACCACTCCAGTGTGTGTGTCCTGCTGTG	562
Db	443	AGAAGTATGAGCTATACGCGAGGACAGACCAACGCGCTGTGTGTGTCCTGCTGTG	502
Qy	563	AGAGTCAACCAAGAGATGAGGAGGAGAGAGAG 597	
Db	503	AGAGTCAACCAAGAGGAGGAGGAGAGAGAG 537	

RESULT	8	AF282255	621 bp	mRNA	linear	ROD 17-NOV-2000
LOCUS	AF282255					
DEFINITION	Mus musculus Down syndrome candidate region 1 protein (Dscr1) mRNA, complete cds.					
ACCESSION	AF282255					
VERSION	AF282255.1					
KEYWORDS	GI:9652249					
SOURCE	Mus musculus.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 621)					
TITLE	Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zamotti, M.					
JOURNAL	The murine DSCR1-like (Down syndrome candidate region 1) gene family: conserved synteny with the human orthologue genes					
PUBMED	Gene 257 (2), 223-232 (2000)					
REFERENCES	20534792					
AUTHORS	2 (bases 1 to 621)					
TITLE	Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zamotti, M.					
JOURNAL	Direct Submission					
FEATURES	Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia Generale, Universite di Bologna, Via Belmeloro, 8, Bologna, BO 40126, Italy					
source	Location/Qualifiers					
gene	1..621					
CDS	/organism="Mus musculus"					
	/strain="BALB/c"					
	/db_xref="taxon:10090"					
	/chromosome="16"					
	/sex="male"					
	/tissue_type="brain"					
	1..621					
	/gene="Dscr1"					
	8..604					
	/gene="Dscr1"					
	/note="similar to the Homo sapiens Down syndrome candidate region 1 protein"					
	/codon_start=1					
	/product="Down syndrome candidate region 1 protein"					
	/protein_id="AA091461.1"					
	/db_xref="GI:9652250"					
	/translation="MEVDLIDLPSATLACHIDPRFVVDGLCRKAFESLFTYDKDTTFOYFSPKRVIRINPISADARLRHKTEIGKMLYPAQTLHGSSHLAPNPDKGLDLSPPASPPVCKMKNVDATPVINVDLVAISLIGGEKVELHAATDTTPSVVHVCSDDNEEESEEMERMKRPKRIITGRPEYPIHLS"					
BASE COUNT	166 a 179 c 159 g 117 t					
ORIGIN						
Query Match	71.4%					
Best Local Similarity	93.3%					
Matches	458; Conservative 0; Mismatches 32; Indels 1; Gaps 1,					
Oy	110 AGCTCCCTGATTGCTTGTCGCAAAAGCATGATGCTTCAAGCAAGTAGAGACGGCC 169					
Db	38 AGGCCCAACATCGCTCCACACTGACCCCGCGCTGTTCTGACGGCTGACCGGCGCC 97					
Oy	170 AAATTGTAATCCCTCTTCAGAACTATGACAGGACCACTTCCAGTATTTAAGAGC 229					
Db	98 AAATTGTAATCCCTCTTCAGAACTATGACAGGACCACTTCCAGTATTTAAGAGC 157					
Oy	230 TTCAAAAGTGTCCGGATAACTTACGACACCCCTTATCTCAGCCGATGCGCAGCTGCGG 289					
Db	158 TTCAAAAGTGTCCGGATAACTTACGACACCCCTTATCTGAGGCCGATGCGCAGCTGCGG 217					
Oy	290 CTGCACAGACCGAGTCTCTGGGGAGGAAATGAAGTTGTTATTTTGCTCAGACTTACAC 349					
Db	218 CTGCACAGACCGAGTCTCTGGGGAGGAAATGAAGTTGTTATTTTGCTCAGACTTACAC 277					
Oy	350 ATGAGAACTTACACCTGGCTCCG-CCAATCCGACAAACAGTTCTCATCTTCCCTCCG 408					

Db	278	ATMGAGATTTCACACTGGCTCCGCCCAATCCCGAAACAAGTTCATCTCCCCGCG	337
OY	409	GCCCTCCTCCCGCTGGAACAACTAGAAGATGCCACCCCCTGTATAATTACAT	468
Db	338	GCCCTCCTCCCGCTGGAACAACTAGAAGATGCCACCCCCTGTATAATTACAT	397
OY	469	CTTTTAATATGCCTTCGAAAGCTGGGGCCAGAGAGATGACTCATGACGACGA	528
Db	398	CTTTTAATATGCCTTCGAAAGCTGGGGCCAGAGAGATGACTCATGACGACGA	457
OY	529	GACACACTCCCACTGATGATGATCCAGTGATGAGAGTGACCAAAGATGAGAGAA	588
Db	458	GACACACTCCCACTGATGATGATCCAGTGATGAGAGTGACCAAAGATGAGAGAA	517
OY	589	GAGGAAGAGAT 599	
Db	518	GAGGAAGAGAT 528	
RESULT 9			
AF260717		2125 bp	mRNA linear ROD 22-MAR-2001
LOCUS		Mus musculus Down syndrome candidate region 1 (Dscr1)	mRNA, complete cds.
DEFINITION			
ACCESSION	AF260717		
VERSION	AF260717.1	GI:7839596	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
RECEIVED			



```
Query Match      71.4%; Score 427.8; DB 10; Length 2125;
Best Local Similarity 93.3%; Pred. No. 8e-125;
Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTCCTGTGTGCGAAGCATGATGCTTCAGCGAAGTGAGACAGGCGCC 169
DB 45 AGCGCCACCATCGCTGCGCACCTGACCGCGCGCTGTCGTGGACGGCTGTGCGCGGCC 104
QY 170 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAAGGC 229
DB 105 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAAGGC 164
QY 230 TTCAACGCTGTCGGATAAATTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 289
DB 165 TTCAACGCTGTCGGATAAATTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 224
QY 290 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 349
DB 225 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 284
QY 350 ATAGGAAGTTTCAACCTCGCTCCG - CCAATCCGACAAACAGTTCTCATCTCCCTCCG 408
DB 285 ATAGGAAGTTTCAACCTCGCTCCG - CCAATCCGACAAACAGTTCTCATCTCCCTCCG 344
QY 409 GCCTCTCTCCGTTGGTGGAAACAAGTAGAAGTCCACCCCGTCAATAAATTACGAT 468
DB 345 GCCTCTCTCTCCGTTGGTGGAAACAAGTAGAAGTCCACCCCGTCAATAAATTACGAT 404
QY 469 CTTTATATGCGATCTCCAACTCGGGCCAGGAGAGATGAACTGCATGACGAGACA 528
DB 405 CTTTATATGCGATCTCCAACTCGGGCCAGGAGAGATGAACTGCATGACGAGACA 464
QY 529 GACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 588
DB 465 GACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 524
QY 589 GAGGAAGAGAT 599
DB 525 GAGGAAGAGAT 535

RESULT 10
AX365312
LOCUS      AX365312
DEFINITION Sequence 2 from Patent WO0204491.
ACCESSION AX365312
VERSION    AX365312.1 GI:18697043
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    Williams, S.R. and Rothermel, B.
TITLE      Methods and compositions relating to muscle selective calcineurin
JOURNAL    Patent: WO 0204491-A 2.17-JAN-2002;
Board of Regents, The University of Texas System (US); Williams,
Sanders R. (US); Rothermel, Beverly (US)
FEATURES
source     Location/Qualifiers
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
1..597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23808.1"
/db_xref="GI:18697044"
/translation="NEEVDLQDPSATTACHLDRFVDFGLCRAPKPSLFRPYDKDTT
FOYFKSFRVRINPSNPLAARLRLHKTFLGKEMKLYPAQTLHIGSHLAPND
KQFLISPASPPVGNKVEDATPVINTDLYAISKLGPEKYLHAATDPTPSVVVHV
CESDQNEEEEEEMRKRPKPKIQTTRRPEVTFILS"

BASE COUNT      160 a      170 c      151 g      116 t
ORIGIN
Query Match      71.2%; Score 426.2; DB 6; Length 597;
Best Local Similarity 93.1%; Pred. No. 2e-124;
Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTCCTGTGTGCGAAGCATGATGCTTCAGCGAAGTGAGACAGGCGCC 169
DB 31 AGCGCCACCATCGCTGCGCACCTGACCGCGCGTGTGTCGTGGACGGCTGTGCGCGGCC 90
QY 170 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAAGGC 229
DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAAGGC 150
QY 230 TTCAACGCTGTCGGATAAATTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 289
DB 151 TTCAACGCTGTCGGATAAATTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 210
QY 290 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAAAGTTGTTATTTGCTCAGACTTTACAC 349
DB 211 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAAAGTTGTTATTTGCTCAGACTTTACAC 270
QY 350 ATAGGAAGTTTCAACCTCGCTCCG - CCAATCCGACAAACAGTTCTCATCTCCCTCCG 408
DB 271 ATAGGAAGTTTCAACCTCGCTCCG - CCAATCCGACAAACAGTTCTCATCTCCCTCCG 330
QY 409 GCCTCTCTCCGTTGGTGGAAACAAGTAGAAGTCCACCCCGTCAATAAATTACGAT 468
DB 331 GCCTCTCTCTCCGTTGGTGGAAACAAGTAGAAGTCCACCCCGTCAATAAATTACGAT 390
QY 469 CTTTATATGCGATCTCCAACTCGGGCCAGGAGAGATGAACTGCATGACGAGACA 528
DB 391 CTTTATATGCGATCTCCAACTCGGGCCAGGAGAGATGAACTGCATGACGAGACA 450
QY 529 GACACCACTCCAGTGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 588
DB 451 GACCCCACTCCAGTGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 510
QY 589 GAGGAAGAGAT 599
DB 511 GAGGAAGAGAT 521

RESULT 11
AF237789
LOCUS      AF237789
DEFINITION Mus musculus myocyte-enriched calcineurin interactin protein 1
splice variant 1 mRNA, complete cds.
ACCESSION AF237789
VERSION    AF237789.1 GI:7542525
KEYWORDS
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and
Williams, R.S.
TITLE      A Protein Encoded within the Down Syndrome Critical Region is
Enriched in Striated Muscles and Inhibits Calcineurin Signaling
JOURNAL    J. Biol. Chem. (2000) In press
REFERENCE   2 (bases 1 to 597)
AUTHORS    Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and
Williams, R.S.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-2000) Internal Medicine, University of Texas
Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX
75390, USA
FEATURES
source     Location/Qualifiers
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
```

## CDS

/chromosome="16"  
 1..597  
 /note="similar to Homo sapiens DSCR1; NCIP1; similar to  
 EST sequence of Genbank Accession Number AA734360"  
 /codon\_start=1  
 /product="myocyte-enriched calcineurin interactin protein  
 1 splice variant 1"  
 /protein\_id="AA63485.1"  
 /db\_xref="GI:7542526"  
 /translation="MEVDLDLPSPATIAHLDPVFDLCAKKESEFLRYDKDTT  
 FQYKSPKRVINFSPLAADARLKHTEFLKEMLYFAQTLHGSSHLAPNPD  
 KQPLSPSPVGMKQVEDATPVINLDLYAISKLGPGEYELHAATDTPSVVHV  
 CESDQENBEEBEMERKRPKIKIQRREPYTHLS"

## BASE COUNT

160 a 170 c 151 g 116 t

## ORIGIN

Query Match 71.2% Score 426.2; DB 10; Length 597;  
 Best Local Similarity 93.1%; Pred. No. 2e-124;  
 Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

OY 110 AGCTCCCTGATTCCTGTCGCAACGATGATGCTTCAGGAAAGTGAACCGAGGCC 169  
 DB 31 AGCGCACCATGCGCTCGACCTGACCGCGCGCTGTCGACGCGCTGCGGCGC 90  
 OY 170 AATTGATCCCTCTTCGACATATGACAGACACACCTTCAGATTTTAAAGAC 229  
 DB 91 AATTGATCCCTCTTCGACATATGACAGACACACCTTCAGATTTTAAAGAC 150  
 OY 230 TTCAACGTGTCGCGATTAACCTTCAGCAACCCCTTATCTGACCGGATGCCAGCTGCGG 289  
 DB 151 TTCAACGTGTCGCGATTAACCTTCAGCAACCCCTTATCTGACCGGATGCCAGCTGCGG 210  
 OY 290 CTGACACAGACCGAGTTCTGGGAGAAATGAATTTGATTTTCTGACGCTTACAC 349  
 DB 211 CTGACACAGACCGAGTTCTGGGAGAAATGAATTTGATTTTCTGACGCTTACAC 270  
 OY 350 ATAGGAAGTTCAACCTGCGTCCG-CCAAATCCGCAACAGTTCTCTCCCTCCG 408  
 DB 271 ATAGGAAGTTCAACCTGCGTCCG-CCAAATCCGCAACAGTTCTCTCCCTCCG 330  
 OY 409 GCCTCTCTCCCTGTCGCAACGATGAAGATGCCACCCCTCTCTAAATTACAT 468  
 DB 331 GCCTCTCTCCCTGTCGCAACGATGAAGATGCCACCCCTCTCTAAATTACAT 390  
 OY 469 CTTTATATGCCATCTTCCAAGCTGGGCGCCAGAGAGAAAGTATGATGATGAGGACA 528  
 DB 391 CTTTATATGCCATCTTCCAAGCTGGGCGCCAGAGAGAAAGTATGATGATGAGGACA 450  
 OY 529 GACACCACTCCAGTGTGTGTCACAGTGTGAGAGTACCAAGAGATGAGAGAA 588  
 DB 451 GACACCACTCCAGTGTGTGTCACAGTGTGAGAGTACCAAGAGATGAGAGAA 510  
 OY 589 GAGGAGAGAGAT 599  
 DB 511 GAGGAGAGAGAT 521

RESULT 12  
 AF263240 2224 bp mRNA linear ROD 30-MAY-2000  
 LOCUS AF263240  
 DEFINITION Mus musculus calcineurin inhibitor mRNA, complete cds,  
 alternatively spliced.

ACCESSION AF263240.1 GI:8102013  
 VERSION AF263240.1  
 KEYWORDS  
 SOURCE Mus musculus  
 ORGANISM Mus musculus

REFERENCE  
 1 (bases 1 to 2224)  
 Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.  
 Down syndrome candidate region 1 (Dscr1), one of three  
 alternatively spliced exon 1 transcripts

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2224)  
 AUTHORS Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,  
 Avia, Castelldefels Km. 2.7, L'Hospital de Llobregat, Barcelona  
 08907, Spain

FEATURES  
 source location/Qualifiers  
 1..2224  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"

## CDS

/note="Down syndrome candidate region 1; Dscr1;  
 alternatively spliced exon 1"  
 /codon\_start=1  
 /product="calcineurin inhibitor"  
 /protein\_id="AA72702.1"  
 /db\_xref="GI:8102014"  
 /translation="MKLYFAOTLHGSSHLAPNPDKQPLSPSPVGMKQVEDAT  
 PVINLDLYAISKLGPGEYELHAATDTPSVVHVCSDDQENBEEBEMERKRPK  
 KIQRREPYTHLS"

## BASE COUNT

519 a 554 c 586 g 562 t 3 others

## ORIGIN

Query Match 71.0% Score 425.4; DB 10; Length 2224;  
 Best Local Similarity 99.5%; Pred. No. 4.7e-124;  
 Matches 437; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 162 CCAGGCGCAATTGATTCCTCTTCGACATATGACAGACACCACTTCAGATTT 221  
 DB 196 CGAGGCGCAATTGATTCCTCTTCGACATATGACAGACACCACTTCAGATTT 255  
 OY 222 TTAAGAGCTTCAACGTGTCGCGATTAACCTTCAGCAACCCCTTATCTGACCGGATGCCA 281  
 DB 256 TTAAGAGCTTCAACGTGTCGCGATTAACCTTCAGCAACCCCTTATCTGACCGGATGCCA 315  
 OY 282 GGTGCGGCTGCACAGACCGAGTTCTGGGAGAAATGAAGTGTATTTTCTGACA 341  
 DB 316 GGTGCGGCTGCACAGACCGAGTTCTGGGAGAAATGAAGTGTATTTTCTGACA 375  
 OY 342 CTTTACATATGGAAGTTCAACCTGCGTCCG-CCAAATCCGCAACAGTTCTCTCACT 400  
 DB 376 CTTTACATATGGAAGTTCAACCTGCGTCCG-CCAAATCCGCAACAGTTCTCTCACT 435  
 OY 401 CCCCTCCGCGCTCTCTCCGTTGCTGGAACAAATGAAGATGCCACCCCGTCAATAA 460  
 DB 436 CCCCTCCGCGCTCTCTCCGTTGCTGGAACAAATGAAGATGCCACCCCGTCAATAA 495  
 OY 461 ATTACGATCTTTTATATGCCATCTTCCAAGCTGGGCGCCAGAGAGAAAGTATGATGATG 520  
 DB 496 ATTACGATCTTTTATATGCCATCTTCCAAGCTGGGCGCCAGAGAGAAAGTATGATGATG 555  
 OY 521 CAGCGACAGACCACTCCAGTGTGTGTCACAGTGTGTGAGAGTACCAAGAGATG 580  
 DB 556 CAGCGACAGACCACTCCAGTGTGTGTCACAGTGTGTGAGAGTACCAAGAGATG 615  
 OY 581 AGGAGAGAGAGAGAGAT 599  
 DB 616 AGGAGAGAGAGAGAGAT 634

## RESULT 13

AX365321 2331 bp DNA linear PAT 15-FEB-2002  
 LOCUS AX365321  
 DEFINITION Sequence 11 from Patent WO2004491.

ACCESSION AX365321  
 VERSION AX365321.1 GI:18697049  
 KEYWORDS

## SOURCE

ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE\*  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
CDS  
BASE COUNT  
ORIGIN  
Query Match  
Best Local Similarity 85.1%; Score 413.8; DB 6; Length 2331;  
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
QY 39 GAGCAGTCTTCCTTAAGCGTCTGCCCGTGAAGAAAGCAGATGATTTAGGACTTTA 98  
Db 103 GACTCGTGGTCTGTAGCGCTTTCAGTGAAGAAAGCAGATGATTTAGAACTTTA 162  
QY 99 GCTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACGATGATGTTTCTCAGCGAAAGTG 158  
Db 163 ACTACAGTTTGTAGCTCCCTGATTCCTGTGTGGCAACGATGATGTTTCTCAGCGAAAGTG 222  
QY 159 AGACCGGGCCAAATTTGAATTCCTCTTCAGAACATATGACAGGACACCACTTCCAGT 218  
Db 223 AAACCGGGCCAAATTTGAGTCCCTCTTTAGGAGGTATGACAGGACATCACCTTTTCA 282  
QY 219 ATTTTAAAGCTTCAACAGTCTCGGATAAATCTTCAGCAACCTTATCTGACCGGATG 278  
Db 283 ATTTTAAAGCTTCAACAGTCTCGGATAAATCTTCAGCAACCTTATCTGACCGGATG 342  
QY 279 CCAGGCTCGCGCTGCAACAGCGAGTCTCTGGGAGGAGGAATGAAGTGTATTTGCTC 338  
Db 343 CCAGGCTCGAGCTGCAACAGCTGAGTCTTCTGGAAAGGAATGAAGTGTATTTGCTC 402  
QY 339 AGACTTTACATAGGAAGTTCACACTGCTCGGC- AATCCCGCAACACAGTTCCTCA 397  
Db 403 AGACTTTACATAGGAAGTTCACACTGCTCGGC- AATCCCGCAACACAGTTCCTCA 462  
QY 398 TCTCCCTCCGGCTCTCTCCCGTGGCTGGAAACAGTAGAAGATGCCACCCCGTCA 457  
Db 463 TCTCCCTCCGGCTCTCTCCCGTGGCTGGAAACAGTAGAAGATGCCACCCCGTCA 522  
QY 458 TAAATACGATCTTTTATATGATCTCCAGCTGCGGCGAGAGAGATGATGATCTGC 517  
Db 523 TAAATACGATCTTTTATATGATCTCCAGCTGCGGCGAGAGAGATGATGATGATG 582  
QY 518 ATGACGGGACAGACACACTCCAGTGTGGTGTCCAGCTGTGTGAGAGTGCACCAAGAGA 577  
Db 583 ACAGCGAGTGCACACACTCCAGGCGTGTGTGTGATGATGATGATGATGATGATGATG 642  
QY 578 ATGACGAGGAGGAGAA 594  
Db 643 AGGAGGAGGAGGAGAA 659  
RESULT 14  
LOCUS HSU85267  
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,  
alternative exon 1, complete cds.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REMARK  
COMMENT  
FEATURES  
source  
gene  
CDS  
BASE COUNT  
ORIGIN  
Query Match  
Best Local Similarity 85.1%; Score 413.8; DB 9; Length 2346;  
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
QY 39 GAGCAGTCTTCCTTAAGCGTCTGCCCGTGAAGAAAGCAGATGATTTAGGACTTTA 98  
Db 103 GACTCGTGGTCTGTAGCGCTTTCAGTGAAGAAAGCAGATGATTTAGAACTTTA 162  
QY 99 GCTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACGATGATGTTTCTCAGCGAAAGTG 158  
Db 163 ACTACAGTTTGTAGCTCCCTGATTCCTGTGTGGCAACGATGATGTTTCTCAGCGAAAGTG 222  
QY 159 AGACCGGGCCAAATTTGAATTCCTCTTCAGAACATATGACAGGACACCACTTCCAGT 218  
Db 223 AAACCGGGCCAAATTTGAGTCCCTCTTTAGGAGGTATGACAGGACATCACCTTTTCA 282  
U85267  
U85267.2 GI:7596915  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and  
Estivill,X.  
A new human gene from the Down syndrome critical region encodes a  
proline-rich protein highly expressed in fetal brain and heart  
Hum. Mol. Genet. 4 (10), 1935-1944 (1995)  
96121593  
8595418  
2 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M.A. and Estivill,X.  
Genomic organization, alternative splicing, and expression patterns  
of the DSCR1 (Down syndrome candidate region 1) gene  
Genomics 44 (3), 358-361 (1997)  
97468152  
9325060  
3 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M. and Estivill,X.  
Direct Submission  
Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca  
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona  
08907, Spain  
4 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M. and Estivill,X.  
Direct Submission  
Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca  
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona  
08907, Spain  
Sequence update by submitter  
On Apr 19, 2000 this sequence version replaced gi:2612867.  
Location/Qualifiers  
1..2346  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1-q22.2"  
1..2346  
/gene="DSCR1"  
/gene="DSCR1"  
/note="down syndrome candidate region 1; one of four  
alternatively spliced exon 1"  
/codon\_start=1  
/protein\_id="AAB84372.1"  
/db\_xref="GI:2612868"  
/translation="MHPRNFYSFSLIACVANSDFSESETRAKFESLPTDYKDI  
FOYFKSPKRVIRNPNPFAADARLQHKTEFLGKMKLYFAOTLHGSSHLAPNPD  
KQFLISPPASPPVGMKQVEDATPVINYDLIYALSKLPGKEKYLHAATDTTSPVVHV  
CESDQEKKEEEMRRPKKIOTRRPEYPIHLIS"  
645 a 470 c 547 g 684 t

```

QY 219 ATTTAAGAGCTTCAACGTCGTCCGATAACTTCAGCAACCCCTTATCTGACAGCATG 278
Db 283 ATTTAAGAGCTTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGCATG 342
QY 279 CCAGGCTCGGCTGCAACAAGCCGAGTTCCTGGGGAAGAAATGAGTTGTTGCTC 338
Db 343 CCAGGCTCGAGTCGATTAAGACTGAGTTCTGGGAAAGAAATGAGTTATTTTGTCTC 402
QY 339 AGACTTTACATAGAAAGTTCAACCTGCTCCGCC-ATTCGCCAACAACGTTCTCTCA 397
Db 403 AGACTTTACATAGAAAGTTCAACCTGCTCCGCCAATTCAGACAGAGTTTCTGA 462
QY 398 TCTCCCTTCGGCTCTCTCTCCGTTGGCTGGAACAAGTGAAGATGCAACCCCGCTCA 457
Db 463 TCTCCCTTCGGCTCTCTCTCCGCTGAGGATGGAACAGTGAAGATGCAACCCCGCTCA 522
QY 458 TAAATTAAGATCTTTATATGTCATCTCCAGCTGAGGCAAGAGAGATATGAATGTC 517
Db 523 TAAATTAAGATCTTTATATGTCATCTCCAGCTGAGGCAAGAGAGATATGAATGTC 582
QY 518 ATGACAGCAACAGACCACTCCAGTGTGTGTCTCACTGTGTGAGATGACCAAGAGA 577
Db 583 ACAGAGCGACTGACACCACTCCAGGCTGTGTGTCTCATGTATGTAGATGATCAAGAGA 642
QY 578 ATGAGAGGAAGAGGAA 594
Db 643 AGAGAGGAAGAGAGGAA 659

```

## RESULT 15

```

AX281651 2348 bp DNA linear PAT 02-NOV-2001
LOCUS AX281651
DEFINITION Sequence 60 from Patent WO0177389.
ACCESSION AX281651
VERSION AX281651.1 GI:16608902
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

## REFERENCE 1

```

AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 60 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source 1. 2348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 042176.5"

```

```

BASE COUNT 638 a 473 c 550 g 687 t
ORIGIN

```

```

Query Match 67.6%; Score 404.8; DB 6; Length 2348;
Best Local Similarity 84.1%; Pred. No. 1.8e-117;
Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

```

```

QY 39 GAGGAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGACTTTA 98
Db 110 GAGGAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGAACTTTA 169
QY 99 GCTTCAATTTTAAAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGAAAGTG 158
Db 170 ACTACAGTTTAAAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGAAAGTG 229
QY 159 AGACGAGGCAATTTTAAAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGAAAGTG 218
Db 230 AAACGAGGCAATTTTAAAGCTCTTCTTAAAGCTCTGCCCGTAAAAAGCAAGATGATTTTAGGAAAGTG 289
QY 219 ATTTAAGAGCTTCAACGTCGTCCGATAACTTCAGCAACCCCTTATCTGACAGCATG 278

```

```

Db 290 ATTTAAGAGCTTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGCATG 349
QY 279 CCAGGCTCGGCTGCAACAAGCCGAGTTCCTGGGGAAGAAATGAGTTGTTGCTC 338
Db 350 CCAGGCTCGAGTCGATTAAGACTGAGTTCTGGGAAAGAAATGAGTTATTTTGTCTC 409
QY 339 AGACTTTACATAGAAAGTTCAACCTGCTCCGCCAATTCAGACAGAGTTTCTGA 469
Db 410 AGACTTTACATAGAAAGTTCAACCTGCTCCGCCAATTCAGACAGAGTTTCTGA 469
QY 397 ATCTCCCTTCGGCTCTCTCTCCGTTGGCTGGAACAAGTGAAGATGCAACCCCGCTCA 456
Db 470 ATCTCCCTTCGGCTCTCTCTCCGCTGAGGATGGAACAGTGAAGATGCAACCCCGCTCA 529
QY 457 ATAAATTAAGATCTTTATATGTCATCTCCAGCTGAGGCAAGAGAGATATGAATGTC 516
Db 530 ATAAATTAAGATCTTTATATGTCATCTCCAGCTGAGGCAAGAGAGATATGAATGTC 589
QY 517 CATGACGAGACAGACCACTCCAGTGTGTGTCTCACTGTGTGAGATGACCAAGAG 576
Db 590 CACGACGAGACAGACCACTCCAGGCTGTGTGTCTCATGTATGTAGATGATCAAGAG 649
QY 577 AATGAGAGGAAGAGGAA 594
Db 650 AAGAGAGGAAGAGGAA 667

```

```

Search completed: December 14, 2002, 17:05:13
Job time : 1084.63 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:06:10 ; Search time 664.555 Seconds  
(without alignments)  
14597.874 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gaggtcaagggaacctcca.....gaggaggaggaagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599	100.0	599	9	AA200984
2	567.2	94.7	710	10	BB617325
3	551	92.0	753	13	B1148584
4	549	91.7	980	12	BG296537
5	539.4	90.1	665	13	B1219142
6	533.8	89.1	827	13	B1102432

7	529	88.3	950	12	BF782633
8	510.8	85.3	1158	14	BQ895806
9	477	79.6	501	12	BF385073
10	461.8	77.1	711	13	BG915568
11	427.8	71.4	520	14	BQ749142
12	423	70.6	1507	11	AK010696
13	415.4	69.3	837	9	AUI24628
14	413.8	69.1	718	9	AUI31040
15	413.8	69.1	885	9	AL544755
16	413.8	69.1	931	9	AL543576
17	413.8	69.1	939	9	AL546617
18	413.8	69.1	946	14	BQ278576
19	413.8	69.1	1041	13	BM450020
20	407.8	68.1	710	12	BG475986
21	405.8	67.7	967	14	BQ946642
22	403.8	67.4	740	13	BI463566
23	403.4	67.3	811	12	BG570239
24	402.8	67.2	662	13	BI464521
25	402.8	67.2	751	9	AL576189
26	399	66.6	906	9	AL544313
27	398.6	66.5	939	9	AL554686
28	398	66.4	1078	13	BM541636
29	386	64.4	1047	12	BG287042
30	385.4	64.3	986	12	BG828534
31	384	64.1	769	13	BI767955
32	367.8	61.4	535	14	BM832622
33	367.4	61.3	627	13	BI859506
34	364.4	60.8	525	9	AL037245
35	363.8	60.7	859	12	BF203629
36	359.4	60.0	822	12	BE895008
37	359	59.9	727	13	BI327875
38	349.4	58.3	705	9	AL550372
39	349.2	58.3	867	14	BQ427531
40	349.2	58.3	890	9	AL538796
41	348.8	58.2	828	9	AL536447
42	348.4	58.2	923	9	AL551657
43	348.2	58.1	956	9	AL556803
44	343.8	57.4	1014	9	AL559594
45	341.8	57.1	526	12	BF041330

#### ALIGNMENTS

RESULT 1  
AA200984  
LOCUS mul2ell.r1 Soares thymus\_2NbMT Mus musculus CDNA clone IMAGE:639212  
DEFINITION 5' similar to TR:G1125052 G1125052 DOWN SYNDROME CRITICAL REGION  
PROTEIN 1. ; mRNA sequence.  
ACCESSION AA200984  
VERSION AA200984.1 GI:1795994  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 599)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mousees@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the







RESULT 4  
BG296537  
LOCUS  
DEFINITION  
60239441P1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:4506130 5',  
mRNA sequence.  
BG296537  
ACCESSION  
BG296537.1 GI:13059271  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 980)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LLM10381 row: a column: 11  
High quality sequence start: 3  
High quality sequence stop: 747.  
Location/Qualifiers  
1..980  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_1db="NIH\_MGC\_94"  
/clone\_1lb="NIH\_MGC\_94"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH-MGC Library."

BASE COUNT 265 a 258 c 269 g 188 t

ORIGIN

Query Match 91.7%; Score 549; DB 12; Length 980;  
Best Local Similarity 99.3%; Pred. No. 2.1e-155;  
Matches 593; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

QY 5 TGCAGGAAGACCTCCAGCTGGCTGAGAGAGCGAGTTCGTTAGCGTCTGC 64  
DB 10 TGCAGGAAGACCTCCAGCTGGCTGAGAGAGCGAGTTCGTTAGCGTCTGC 69  
QY 65 CCGGTGAAAAAGAGAGATG-ATTTAGGAGCTTTAGCTACAAATTTTACCTGATTC 123  
DB 70 CCGGTGAAAAAGAGAGATGATTTTAGGAGCTTTAGCTACAAATTTTACCTGATTC 128  
QY 124 TTGTGTGGCAAGATGATGCTTTCAGCGAAAGAGACAGGCGCAATTTGAATCCCT 183  
DB 129 TTGTGTGGCAAGATGATGCTTTCAGCGAAAGAGACAGGCGCAATTTGAATCCCT 188  
QY 184 CTTGAGAACTATGACAGAGACACCTTCAGATTTTAAAGAGCTTCAAGCGTCTCC 243  
DB 189 CTTGAGAACTATGACAGAGACACCTTCAGATTTTAAAGAGCTTCAAGCGTCTCC 248  
QY 244 GATTAACCTTGAAGACCCCTTATCTGAGCCGATGCCAGCTGCGGTGCAAGACCGA 303  
DB 249 GATTAACCTTGAAGACCCCTTATCTGAGCCGATGCCAGCTGCGGTGCAAGACCGA 308  
QY 304 GTTCTGTGGAGAAATGAGTTGTATTTTGTCTACAGCTTACATAGGAAGTTTACA 363  
DB 309 GTTCTGTGGAGAAATGAGTTGTATTTTGTCTACAGCTTACATAGGAAGTTTACA 367

QY 364 CTTGCTCCG-CCAAATCCGACAAAGATTCTCATCTCCCTCCGCGCTCTCCCGT 422  
DB 368 CTTGCTCCGCGCCCAATCCGACAAAGATTCTCATCTCCCTCCGCGCTCTCCCGT 427  
QY 423 TGGTGGAAACAGTAAAGATGCCACCCCGCATTAATTTAGATCTTTATATGCGAT 482  
DB 428 TGGTGGAAACAGTAAAGATGCCACCCCGCATTAATTTAGATCTTTATATGCGAT 487  
QY 483 CTCGAAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 542  
DB 488 CTCGAAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 547  
QY 543 TGTGTGTGTCACGTGTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 599  
DB 548 TGTGTGTGTCACGTGTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 604

RESULT 5  
B1219142  
LOCUS  
DEFINITION  
602935788F1 NCI\_CGAP\_L19 Mus musculus cDNA clone IMAGE:5093372 5',  
mRNA sequence.  
B1219142  
ACCESSION  
B1219142.1 GI:14672586  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 665)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LLM11238 row: o column: 21  
High quality sequence stop: 657.  
Location/Qualifiers  
1..665  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone\_1db="NCI\_CGAP\_L19"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: liver; Vector: pCMV-Sport6; Site: 1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.9 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 192 a 170 c 158 g 145 t

ORIGIN

Query Match 90.1%; Score 539.4; DB 13; Length 665;  
Best Local Similarity 97.6%; Pred. No. 1.4e-157;  
Matches 579; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 9 AAGGAACCTCCAGCTTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 68  
DB 5 AAGGAACCTCCAGCTTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64  
QY 69 TGAAGAAAGAGATG-ATTTAGGAGCTTTAGCTACAAATTTTACCTGATTCCTGT 127  
DB 65 TGAAGAAAGAGATGATTTTAGGAGCTTTAGCTACAAATTTTACCTGATTCCTGT 124  
QY 128 GTGGCAAGAGATGATCTTCAAGCGAAAGTGAAGACAGGCGCAATTTGAATCCCTCTTC 187



```

Db 125 GTGCAACGATGATGTTCTTCAGCAAGTGAGACCA-GGCCAAATTTGAATCCCTCTTC 183
Qy 188 AGAACAATATGACAGGACACCACTTCCAGTATTTTAAGAGCTTCAAAACGTGTCCGGATA 247
Db 184 AGAACAATATGACAGGACACCACTTCCAGTATTTTAAGAGCTTCAAAACGTGTCCGGATA 243
Qy 248 AACTTCAGCAACCCCTTATCTGACGCGATGCCAGGCTGGGCTGCACAAGACCGAGTTC 307
Db 244 AACTTCAGCAACCCCTTATCTGACGCGATGCCAGGCTGGGCTGCACAAGACCGAGTTC 303
Qy 308 CTGGGGAAGGAAATGAAGTTGTATTTTCTCAGACTTTTACATAGGAAGTTCCACCTG 367
Db 304 CTGGGGAAGGAAATGAAGTTGTATTTTCTCAGACTTTTACATAGGAAGTTCCACCTG 363
Qy 368 GCTCCG-CCAAATCCGCAACAGTTCCTCATCTCCCTCCGCTCTCTCCCGTTGGC 426
Db 364 GCTCCGCCCAATCCGCAACAGTTCCTCATCTCCCTCCGCTCTCTCCCGTTGGC 423
Qy 427 TGGAAACAAGTAGACCAATTTATACCCCGTCATAAATTTACGATCTTTTATATGCCATCTCC 483
Db 424 TGGAAACAAGTAGACCAATTTATACCCCGTCATAAATTTACGATCTTTTATATGCCATCTCC 483
Qy 487 AAGCTGGGGCCAGGAGAGAGTATGAATCTGCATGCGAGCAGACACCACTCCAGTGTG 546
Db 484 AAGCTGGGGCCAGGAGAGAGTATGAATCTGCATGCGAGCAGACACCACTCCAGTGTG 543
Qy 547 GTGCTCACGTGTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGGAAGAGAT 599
Db 544 GTGCTCACGTGTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGGAAGAGAT 596

RESULT 6
BI102432
LOCUS 602889619F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044612
DEFINITION 5', mRNA sequence.
ACCESSION BI102432
VERSION BI102432.1 GI:14553325
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1122 row: f column: 05
High quality sequence stop: 666.
Location/Qualifiers
1. .827
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5044612"
/lab_host="NCI CGAP_Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library. |"
206 a 248 c 204 g 169 t

BASE COUNT
ORIGIN

```

```

Query Match 89.1%; Score 533.8; DB 13; Length 827;
Best Local Similarity 98.8%; Pred. No. 7.8e-151;
Matches 590; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 6 GCAAGGAACCTCCAGCTTGGCTTGGCTGAGAGAGCGAGTGGTTCGTTAAAGCGTCTGCC 65
Db 1 GCAAGGAACCTCCAGCTTGGCTTGGCTGAGAGAGCGAGTGGTTCGTTAAAGCGTCTGCC 60
Qy 66 CCGTGAANAAGCAGATG-ATTTAGGACCTTTAGTACAATTTTAGCTCCCTGATTGT 124
Db 61 CCGTGAANAAGCAGATGCAATTTTAGGACCTTTAGTACAATTTTAGCTCCCTGATTGT 120
Qy 125 TGTGTGGCAAAACGATGATCTTTCAGCGAAAGTGAGACCAGGCCCAAAATTTCAATCCCTC 184
Db 121 TGTGTGGCAAAACGATGATCTTTCAGCGAAAGTGAGACCA-GGCCAAATTTGAATCCCTC 179
Qy 185 TTCAGAACATATGACAGGACACCACTTTCAGTATTTTAAGAGCTTCAACGTTCCCG 244
Db 180 TTCAGAACATATGACAGGACACCACTTTCAGTATTTTAAGAGCTTCAACGTTCCCG 239
Qy 245 ATAACTTCAGCAACCCCTTATCTGACGCGATGCCAGGCTGGGCTGCACAAGACCGAG 304
Db 240 ATAACTTCAGCAACCCCTTATCTGACGCGATGCCAGGCTGGGCTGCACAAGACCGAG 299
Qy 305 TTCTCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCCACAC 364
Db 300 TTCTCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCCACAC 359
Qy 365 CTGGCTCCG-CCAAATCCCGCAAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTCCCGTT 423
Db 360 CTGGCTCCGCCCCAATCCCGCAAAACAGTTCCTCATCTCCCTCCGCTCTCTCCCTCCCGTT 419
Qy 424 GGCTGAAACAGTAGAAGATGCCACCCCGTCATAAATTTAGCATCTTTTATATGCCATC 483
Db 420 GGCTGAAACAGTAGAAGATGCCACCCCGTCATAAATTTAGCATCTTTTATATGCCATC 479
Qy 484 TCCAAGCTGGGCGCCAGGAGAGAGTATGAATCTGCATGCGAGCAGACACCACTCCAGT 543
Db 480 TCCAAGCTGGGCGCCAGGAGAGAGTATGAATCTGCATGCGAGCAGACACCA-TCCTAGT 538
Qy 544 GTGGTGGTCCAGCTGTGTGAGAGTGACCAAGAGATAGG-AGGAAGAGGAGAGAT 599
Db 539 GTGGTGGTCCAGCTGTGTGAGAGTGACCCAGAGATGAGGAGGAGGAGGAGAGAT 595

RESULT 7
BI102432
LOCUS 602107462F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:4236038
DEFINITION 5', mRNA sequence.
ACCESSION BI102432
VERSION BI102432.1 GI:12087669
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9843 row: k column: 15
High quality sequence stop: 701.

```



BF385073 501 bp mRNA linear EST 27-NOV-2000  
 LOCUS 602045671P1 NCI\_CGAP\_Li9 Mus musculus cDNA clone IMAGE:4195384 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF385073  
 VERSION BF385073.1 GI:11366365  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 501)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9530 row: m column: 17  
 High quality sequence stop: 501.  
 Location/Qualifiers  
 1..501  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4195384"  
 /clone\_lib="NCI\_CGAP\_Li9"  
 /lab\_host="DHI08 (T1 phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.9 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 132 a 134 c 113 g 122 t  
 ORIGIN  
 Query Match 79.6%; Score 477; DB 12; Length 501;  
 Best Local Similarity 99.6%; Pred. No. 1e-133;  
 Matches 499; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
 QY 16 CTCAGCTGGGCTTACTCAGAGAGCGAGTCGTTGCTTAAGCGTGTGCCCGTGAAGAA 75  
 Db 1 CTCAGCTGGGCTTACTCAGAGAGCGAGTCGTTGCTTAAGCGTGTGCCCGTGAAGAA 60  
 QY 76 GCAGATG-ATTTAGGACCTTAGCTACAAATTTAGCTCCCTGATGCTTGTGTGGCAA 134  
 Db 61 GCAGATGCAATTTAGGACCTTTAGCTACAAATTTAGCTCCCTGATGCTTGTGTGGCAA 120  
 QY 135 ACATGATGCTTTCAGCGAAGTGAGACGAGCGCCAAATTTGAATCCCTCTTCAGAACAT 194  
 Db 121 ACATGATGCTTTCAGCGAAGTGAGACGAGCGCCAAATTTGAATCCCTCTTCAGAACAT 180  
 QY 195 ATCAAGAGACACACCTTCCAGTATTTAAGAGCTTCAAACTGTCGGATAAACTTCA 254  
 Db 181 ATCAAGAGACACACCTTCCAGTATTTAAGAGCTTCAAACTGTCGGATAAACTTCA 240  
 QY 255 GCAACCCCTTATCTGAGCGGATGCGAGGTCGAGCGTGCAGAGACCGAGTTCCTGGGGA 314  
 Db 241 GCAACCCCTTATCTGAGCGGATGCGAGGTCGAGCGTGCAGAGACCGAGTTCCTGGGGA 300  
 QY 315 AGGAATGAAGTTGATTTTGTCTCAGACTTTACACATAGGAAAGTTTACACCTGGCTCCG- 373  
 Db 301 AGGAATGAAGTTGATTTTGTCTCAGACTTTACACATAGGAAAGTTTACACCTGGCTCCG 360  
 QY 374 CCAATCCGACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAC 433  
 Db 361 CCAATCCGACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAC 420

QY 434 AACTAGAGATGCCACCCCGTCATAAATTCAGATCTTTTATATGCGATCTCCAACTGG 493  
 Db 421 AACTAGAGATGCCACCCCGTCATAAATTCAGATCTTTTATATGCGATCTCCAACTGG 480  
 QY 494 GGCAGGAGAGAGTAGTATGAAC 514  
 Db 481 GGCAGGAGAGAGTAGTATGAAC 501  
 RESULT 10  
 BG915568  
 LOCUS 602815815F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4945346 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG915568  
 VERSION BG915568.1 GI:14296044  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 711)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
 Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM10891 row: n column: 03  
 High quality sequence stop: 711.  
 Location/Qualifiers  
 1..711  
 /organism="Mus musculus"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4945346"  
 /clone\_lib="NCI\_CGAP\_Mam4"  
 /tissue\_type="tumor\_gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DHI08"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Priscilla Furth,  
 NIH Reference for transgenic model: Li et al., Cell Growth  
 and Differentiation 7, 3-11 (1996)."  
 BASE COUNT 196 a 184 c 183 g 148 t  
 ORIGIN  
 Query Match 77.1%; Score 461.8; DB 13; Length 711;  
 Best Local Similarity 99.0%; Pred. No. 5.3e-129;  
 Matches 496; Conservative 0; Mismatches 2; Indels 3; Gaps 3;  
 QY 101 TACAATTTAGCTCCCTGATGCTTGTGTGGCAACCATGATGCTTTCAGGAAAGTGAG 160  
 Db 1 TACAATTTAGCTCCCTGATGCTTGTGTGGCAACCATGATGCTTTCAGGAAAGTGAG 60  
 QY 161 ACCAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGGACACCACTTCCAGTAT 220  
 Db 61 ACCA-GGCCAAATTTGAATCCCTCTTCAGAACATATGACAGGACACCACTTCCAGTAT 119  
 QY 221 TTTAAGAGCTTCAAACTGTCCGATAAACTTCAGCAACCCCTTATCTGAGCGGATGCC 280  
 Db 120 TTTAAGAGCTTCAAACTGTCCGATAAACTTCAGCAACCCCTTATCTGAGCGGATGCC 179  
 QY 281 AGGCTGGGCTGCACAGACCGAGTTCTCTGGGGAAGGAAATGAAGTTGTATTGCTCAG 340

```

Db 180 AGCGTGGCTCACAAGCCAGTTCCTGGGAGAAATGATGATTTCTCAG 239
Qy 341 ACTTACACATAGGAAGTTACACCTGCTCCG-CCATCCCGACAAACAGTTCTCATC 399
Db 240 ACTTACACATAGGAAGTTACACCTGCTCCGCGCCCAATCCCGAACAACAGTTCTCATC 299
Qy 400 TCCCTCCCGGCTCT-CCCTCCGTTGGCTGGAACAGTAGAAGTGCACCCCGTCAT 458
Db 300 TCCCTCCCGGCTCTCTCCCTCCGCTGGCTGGAACAGTAGAAGTGCACCCCGTCAT 359
Qy 459 AATTACGATCTTTATATATGATCCATCTCAAGCTGGGCGCAGAGAGAAATGAACTGCA 518
Db 360 AATTACGATCTTTATATATGATCCATCTCAAGCTGGGCGCAGAGAGAAATGAACTGCA 419
Qy 519 TCGAGCAGACAGACCACTCTCCAGTGTGTGTGCTCACTGTGTGAGAGACCAAGAGA 578
Db 420 TCGAGCAGACAGACCACTCTCCAGTGTGTGTGCTCACTGTGTGAGAGACCAAGAGA 479
Qy 579 TGAAGAGAGAGAGAGAGAT 599
Db 480 TGAAGAGAGAGAGAGAGAT 500

```

RESULT 11 520 bp mRNA linear EST 17-JUL-2002

LOCUS B0749142 UI-M-FD0-bvd-c-04-0-UI.r1 NIH\_BMAP\_FD0 Mus musculus cDNA clone  
DEFINITION IMAGE:5716539 5', mRNA sequence.

ACCESSION B0749142 GI:21895929

VERSION EST.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 520)

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapds-rc@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 520

/organism="Mus musculus"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:5716539"

/clone\_1lb="NIH BMAP FD0"

/clone\_type="whole Brain"

/dev\_stage="embryo 12.5 dpc"

/lab\_host="DH10B (T1 phage resistant)"

/note="Organ: Brain; Vector: pYX-Asc; Site: 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 137 a 147 c 131 g 105 t

Query Match 71.4%; Score 427.8; DB 14; Length 520;

Best Local Similarity 93.3%; Pred. No. 9.3e-119;

Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

Qy 110 AGCTCCGATGTTGTTGTGTGGAACATGATGTTCTTCGCAAGAGTGAACAGGAGCC 169

Db 12 AGCGCCACATGCGCTTCCACCTGAGCCCGCGGTGCTTGAGAGCGCTGTCCGGGCC 71

Qy 170 AATTGATCCCTCTTCAGAACATATGACAGACACCTTCAGTATTTAAGAGC 229

Db 72 AATTGATCCCTCTTCAGAACATATGACAGACACCTTCAGTATTTAAGAGC 131

Qy 230 TTCAACGTGTCCGATTAACCTTACGACACCCCTTATCTGAGCGCATTCGAGCTGGG 289

Db 132 TTCAACGTGTCCGATTAACCTTACGACACCCCTTATCTGAGCGCATTCGAGCTGGG 191

Qy 290 CTGACAGACCGAGTTCCTGGGAGAAATGAAGTTGATTTTGTCTCAACTTTACAC 349

Db 192 CTGACAGACCGAGTTCCTGGGAGAAATGAAGTTGATTTTGTCTCAACTTTACAC 251

Qy 350 ATAGGAAGTTCAACCTGCTCCG-CCATCCGACAAACAGTTCTCATCTCCCTCCG 408

Db 252 ATAGGAAGTTCAACCTGCTCCG-CCATCCGACAAACAGTTCTCATCTCCCTCCG 311

Qy 409 GCCTTCTCCCGTTGGTGTGGAACAGTAGAATGCAACCCCGTCATTAATGAGAT 468

Db 312 GCCTTCTCCCGTTGGTGTGGAACAGTAGAATGCAACCCCGTCATTAATGAGAT 371

Qy 469 CTTTATATGATCTCCAGCTGGGCGCAGAGAGATGATGATGATGATGATGATGATGAT 528

Db 372 CTTTATATGATCTCCAGCTGGGCGCAGAGAGATGATGATGATGATGATGATGATGAT 431

Qy 529 GACACCACTCCAGTGTGTGTGTCACCTGTGTGAGAGTGCACCAAGAGATGAGAGAA 588

Db 432 GACACCACTCCAGTGTGTGTGTCACCTGTGTGAGAGTGCACCAAGAGATGAGAGAA 491

Qy 589 GAGGAAGAGAT 599

Db 492 GAGGAAGAGAT 502

RESULT 12

LOCUS AK010696 1507 bp mRNA linear HTC 19-JAN-2002

DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,

clone:2410048A02:Down syndrome critical region homolog 1 (human),

full insert sequence.

ACCESSION AK010696 GI:12846329

VERSION HTC, CAP trapper

KEYWORDS Mus musculus (strain: C57BL/6J) ES cells cDNA to mRNA,

clone\_1lb:RIKEN full-length enriched mouse cDNA library

clone:2410048A02.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

REFERENCE Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE 99279253

PUBMED 10349636

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itch, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

#### REFERENCE

#### AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hatada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer

#### TITLE

#### JOURNAL

#### MEDLINE

#### PUBMED

#### REFERENCE

#### AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischnann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barch, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

#### TITLE

#### JOURNAL

#### MEDLINE

#### PUBMED

#### REFERENCE

#### AUTHORS

Adachi, J., Aizawa, K., Akahita, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayashizaki, Y., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tegami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

#### TITLE

#### JOURNAL

#### MEDLINE

#### PUBMED

#### REFERENCE

#### AUTHORS

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTITTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase

and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5',

GAGAGAGAGATTCGCGAGTTAATTAATCCCCCCCCCC 3'], cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

#### FEATURES

#### source

1..1507  
/organism="Mus musculus"  
/strain="CS7BL/6J"  
/db\_xref="PANTOM:DB:2410048A02"  
/db\_xref="MGD:MG1:1898605"  
/db\_xref="taxon:10090"  
/clone="2410048A02"  
/cell\_type="ES cells"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"

#### gene

1..1507

#### CDS

208..564  
/gene="Dscr1"  
/note="Down syndrome critical region homolog 1 (human)"  
data source:MGD, source key:MG1:1890564, evidence:ISS putative"  
/codon\_start=1  
/protein\_id="BAB27128.1"  
/db\_xref="GI:12846330"  
/db\_xref="MGD:MG1:1890564"  
/translation="MKLYFACTLHIGSSHLAPNPDKFLISPPASPPVGMKQVEDAT PVINYLILYAIKSLGPGKYEKELHAATDTTPSVVHVCESDQENEEEMERMKRPK KIIOTREPTPIHLS"  
BASE COUNT 354 a 402 c 390 g 361 t

#### ORIGIN

Query Match 70.6%; Score 423; DB 11; Length 1507;

Best Local Similarity 99.8%; Pred. No. 5.2e-117;

Matches 434; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy	166	GGCCAAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTAA	225
Db	54	GGCCAAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTAA	113
Qy	226	GAGCTTCAAAAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCGCGGATGCCAGCT	285
Db	114	GAGCTTCAAAAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCGCGGATGCCAGCT	173
Qy	286	CGGCTGCACAAAGCCGAGTTCCTGGGGAAGAAATGAAGTTGTTTGTCTCAGACTTT	345
Db	174	CGGCTGCACAAAGCCGAGTTCCTGGGGAAGAAATGAAGTTGTTTGTCTCAGACTTT	233
Qy	346	ACACATAGGAAGTTTCACACCTGGCTCCG-CCAAATCCCGACAAACAGTTCCTCATCTCCC	404
Db	234	ACACATAGGAAGTTTCACACCTGGCTCCG-CCAAATCCCGACAAACAGTTCCTCATCTCCC	293
Qy	405	TCCGGCTCTCTCTCCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTA	464
Db	294	TCCGGCTCTCTCTCCCGTTGGCTGGAAACAAAGTAGAAGATGCCACCCCGTCATAAATTA	353
Qy	465	CGATCTTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGATGATGAATGCACTGCAGC	524
Db	354	CGATCTTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGATGATGAATGCACTGCAGC	413
Qy	525	GACAGACACCACTCCCACTGTGGTGGTCCACGCTGTGTGACAGTGACCAAGAGATGAGGA	584
Db	414	GACAGACACCACTCCCACTGTGGTGGTCCACGCTGTGTGACAGTGACCAAGAGATGAGGA	473
Qy	585	GAAGAGGAAGAGAT	599
Db	474	GAAGAGGAAGAGAT	488

#### RESULT 13

AU124628

LOCUS AU124628 837 bp mRNA linear EST 01-AUG-2002  
 DEFINITION AU124628 NT2RM4 Homo sapiens CDNA clone NT2RM4000307 5', mRNA  
 ACCESSION AU124628  
 VERSION AU124628  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Ota,T., Makamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.  
 HRI human CDNA project (Ota,T., Makamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., and Isogai,T.)  
 Y., Sugano,S., Isogai,T.)  
 Unpublished (2000)  
 CONTACT: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genom@hri.co.jp  
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
 Location/Qualifiers  
 1..837  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="NT2RM4000307"  
 /clone\_lib="NT2RM4"  
 /cell\_type="teratocarcinoma"  
 /note="Vector: PME18FLJ; mRNA from uninduced NT2 neuronal precursor cells"  
 BASE COUNT 241 a 201 c 207 g 186 t 2 others  
 ORIGIN  
 Query Match 69.3%; Score 415.4; DB 9; Length 837;  
 Best Local Similarity 85.3%; Pred. No. 7.4e-115;  
 Matches 475; Conservative 0; Mismatches 81; Indels 1; Gaps 1;  
 39 GAGCGATCGTCTGTAAGCTCTGCCCGTGAAGAAAGCAGATGATTTAGGACTTTA 98  
 77 GACTGCGTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGCAATTTAGAACTTTA 136  
 99 GCTACATTTTAACTCCCTGATTTGCTGTGTGGCAACGATGATGCTTTCAGCGAAAGTG 158  
 137 ACTACAGTTTAACTCCCTGATTTGCTGTGTGGCAACGATGATGCTTTCAGCGAAAGTG 196  
 159 AGACCAAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCAGT 218  
 197 AACACAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACATACCTTTAGT 256  
 219 ATTTAAGAGCTTCAAAAGTGTCCGGATTAACCTTCAGAACCTCTTATCTGACGCGATG 278  
 257 ATTTAAGAGCTTCAAAAGTGTCCGGATTAACCTTCAGAACCTCTTATCTGACGCGATG 316  
 279 CCAGGCTGCGGCTGCACAAAGACCGAGTTCTCTGGGAAAGAAATGAAGTTGATTTTCTC 338  
 317 CCAAGCTGCGGCTGCATAGAGTGTCTGGGAAAGAAATGAAGTTAATTTTCTC 376  
 339 AGACTTACATAGAGAAAGTTTCAACCTGCTCCGCC-AAATCCGACAAACAGTTCTCTCA 397  
 377 AGACTTACATAGAGAAAGTTTCAACCTGCTCCGCC-AAATCCGACAAACAGTTTCTCA 436  
 398 TCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAACAAAGTGAAGATGCCACCCCGTCA 457  
 437 TCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAACAAAGTGAAGATGCCACCCCGTCA 496

RESULT 14  
 AU131040  
 LOCUS AU131040 718 bp mRNA linear EST 01-AUG-2002  
 DEFINITION AU131040 NT2RP3 Homo sapiens CDNA clone NT2RP3001895 5', mRNA  
 ACCESSION AU131040  
 VERSION AU131040  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.  
 HRI human CDNA project  
 Unpublished (2000)  
 CONTACT: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genom@hri.co.jp  
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
 Location/Qualifiers  
 1..718  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="NT2RP3001895"  
 /clone\_lib="NT2RP3"  
 /cell\_type="teratocarcinoma"  
 /note="Vector: PME18FLJ; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"  
 BASE COUNT 217 a 174 c 169 g 155 t 3 others  
 ORIGIN  
 Query Match 69.1%; Score 413.8; DB 9; Length 718;  
 Best Local Similarity 85.1%; Pred. No. 2e-114;  
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
 39 GAGCGATCGTCTGTAAGCTCTGCCCGTGAAGAAAGCAGATGATTTAGGACTTTA 98  
 61 GACTGCGTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGCAATTTAGAACTTTA 120  
 99 GCTACATTTTAACTCCCTGATTTGCTGTGTGGCAACGATGATGCTTTCAGCGAAAGTG 158  
 121 ACTACAGTTTAACTCCCTGATTTGCTGTGTGGCAACGATGATGCTTTCAGCGAAAGTG 180  
 159 AGACCAAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCTTCCAGT 218  
 181 AACACAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACATACCTTTAGT 240  
 219 ATTTAAGAGCTTCAAAAGTGTCCGGATTAACCTTCAGAACCTCTTATCTGACGCGATG 278

```

Db 241 *ATTTTAAGAGCTTCAACAGAGTGAAGTAACTTACAGCAACCCCTTCTCCGACGAGATG 300
QY 279 CAGGCTGGCGGTGCACAGACCGAGTTCCTGGGGAAGGAAATGAAGTTGATTTTGCTC 338
Db 301 CAGGCTCCAGCTGCATAGAGCTGAGTTCTGGGAAGGAAATGAAGTTATATTTTGCTC 360
QY 339 AGACTTTACACATAGAGAGTTACACCTGGCTCCGCGC-AATCCGACAAACAGTTTCTCA 397
Db 361 AGACCTTTACACATAGAGAGTACACCTGGCTCCGCGCAATCCAGCAAGCAGTTTCTGA 420
QY 398 TCTCCCTCCGCGCTCTCCCTCCGCTGGCTGGGAACAGTAGAAGTGCACCCCGCTCA 457
Db 421 TCTCCCTCCGCGCTCTCCGCGAGTGGGATGGAACAAGTGAAGTGCACCCCGCTCA 480
QY 458 TAAATTTACATCTTTTATATGCACTCTCCAAGCTGGGCGCAGGAGAGTATGAAGTGC 517
Db 481 TAAACTATGATCTCTATATGCACTCTCCAAGCTGGGCGCAGGAGAGTATGAAGTGC 540
QY 518 ATGACGAGACAGACCACTCCAGTGGTGGTCCACGTGTGTGAGAGTGACCAAGAGA 577
Db 541 AGCAGCGACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTATCAAGAGA 600
QY 578 ATGAGAGGAGAGGAA 594
Db 601 AGGAGGAGAGAGGAA 617

RESULT 15
AL544755 885 bp mRNA linear EST 16-FEB-2001
LOCUS AL544755 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1022YJ23 5
DEFINITION prime, mRNA sequence.
ACCESSION AL544755
VERSION AL544755.1 GI:12877235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..885
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1022YJ23"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT)-primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 247 a 211 c 228 g 198 t 1 others
ORIGIN

Query Match 69.1%; Score 413.8; DB 9; Length 885;
Best Local Similarity 85.1%; Pred. No. 2.3e-114;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGCGAGTCGTTGTTAAGCGTCTGCCCGTGGAAGAGAGAGTATGATTTTAGGACTTTA 98

```

Search completed: December 14, 2002, 18:10:49  
 Job time : 670.555 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 380.679 Seconds  
(without alignments)  
13789.580 Million cell updates/sec

Title: US-09-782-953-11

Perfect score: 2331

Sequence: 1 tttttttccccaggaggt.....gggcacagtatgctccttac 2331

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	100.0	2331	24	Human MCIP associa
2	2318.4	99.5	2348	24	Human DNA sequence
3	2109	90.5	2212	24	Human MCIP associa
4	2101	90.1	2173	24	Gene #3341 used to
5	2094.2	89.8	2358	21	Lung cancer associ
6	2059.2	88.3	2174	20	Lung DSCR1 coding
7	486	20.8	486	22	Human breast cell
8	486	20.8	486	22	Human foetal liver
9	486	20.8	486	22	Probe #809 for gen

C 10	486	20.8	486	22	AAK00816	Human brain expres
C 11	486	20.8	486	22	AAK26270	Human bone marrow
C 12	486	20.8	486	22	AAI10903	Probe #836 for gen
C 13	486	20.8	486	22	AAI32163	Probe #849 used to
C 14	486	20.8	486	22	AAI00825	Probe #816 used to
C 15	486	20.8	486	22	ABS00858	Human genome-deriv
C 16	470.6	20.2	597	24	AAAD30153	Mouse MCIP associa
C 17	442	19.0	442	21	AAAC01774	Human secreted pro
C 18	413.8	17.8	599	24	AAAD30151	Human MCIP associa
C 19	403.8	17.3	597	24	AAAD30152	Human MCIP associa
C 20	382.4	16.4	404	22	AAAL23689	Human breast tumor
C 21	365.2	15.7	385	22	AAH55546	Human breast tumor
C 22	327.6	14.1	640	24	ABQ59719	Human breast cancer
C 23	246.8	10.6	251	20	AAH84229	DNA encoding human
C 24	246.8	10.6	251	21	AAAC79458	Known human breast
C 25	246.8	10.6	251	24	ABK29002	Human breast tumor
C 26	194.4	8.3	3184	24	AAAD30157	Human MCIP associa
C 27	194.4	8.3	3184	24	ABL61768	Colon adenocarcino
C 28	194.4	8.3	3184	24	ABL65187	Lung cancer relate
C 29	193.2	8.3	594	24	AAAD30154	Human MCIP associa
C 30	189	8.1	720	24	AAAD30159	Human MCIP associa
C 31	189	8.1	828	24	AAAD30158	Human MCIP associa
C 32	187	8.0	934	22	AAF25338	Nucleotide sequenc
C 33	187	8.0	3159	24	ABA91457	Human Down syndrom
C 34	176.2	7.6	415	22	AAAL14827	Human breast cancer
C 35	175	7.5	412	22	ABA56172	Human foetal liver
C 36	175	7.5	412	22	AAK04364	Human brain expres
C 37	175	7.5	412	22	AAI14449	Probe #4382 for ge
C 38	175	7.5	412	22	AAI35821	Probe #4507 used t
C 39	175	7.5	412	22	AAI04272	Probe #4263 used t
C 40	175	7.5	412	24	ABSO4419	Human genome-deriv
C 41	175	7.5	446	22	ABA34406	Human breast cell
C 42	175	7.5	446	22	ABA53852	Human foetal liver
C 43	175	7.5	446	22	ABA23598	Probe #2064 for ge
C 44	175	7.5	446	22	AAK02113	Human brain expres
C 45	175	7.5	446	22	AAK27563	Human bone marrow

## ALIGNMENTS

RESULT 1  
AAD30155  
ID AAD30155 standard; DNA; 2331 BP.

XX AAD30155;

XX 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #1.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 144...737

FT /\*tag= a

FT /product= "Human MCIP associated protein #1"

XX WO200204491-A2.

PD 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.





QY 660 ATGGAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACGAGGCGGAGTACG 719  
 Db 668 ATGGAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACGAGGCGGAGTACG 727  
 QY 720 CCGATCCACCTCAGCTGAACCTGGACGCGACGAGGAGCCATTTCCAAATCATCTCACGG 779  
 Db 728 CCGATCCACCTCAGCTGAACCTGGACGCGACGAGGAGCCATTTCCAAATCATCTCACGG 787  
 QY 780 AGGAATCTTTTACTGTGAGAGTGGCTGGTCAAGACTTCTTGGAGAGTGGACGCGAGATC 839  
 Db 788 AGGAATCTTTTACTGTGAGAGTGGCTGGTCAAGACTTCTTGGAGAGTGGACGCGAGATC 847  
 QY 840 GGGGTGGGCAAGATCCAGTTCATGTGCTCAGAAAGAAATCAAGGCGGTGCCCTTGT 899  
 Db 848 GGGGTGGGCAAGATCCAGTTCATGTGCTCAGAAAGAAATCAAGGCGGTGCCCTTGT 907  
 QY 900 TCTAATGCTGCACACCACTTACTGTTCATGCGACCCGGGAAATGACTTGGGCAATCAGTG 959  
 Db 908 TCTAATGCTGCACACCACTTACTGTTCATGCGACCCGGGAAATGACTTGGGCAATCAGTG 967  
 QY 960 AGTTGTGCTGATCGCACAAAGACATTTGGGACTGTCTTGAAGAAAACAGATAATGATAGT 1019  
 Db 968 AGTTGTGCTGATCGCACAAAGACATTTGGGACTGTCTTGAAGAAAACAGATAATGATAGT 1027  
 QY 1020 GTTTGTACTGTCTTCTTCTGAGTACTGTCTGTCTGTGCAAGGCGGAGGTGATCAGTGA 1079  
 Db 1028 GTTTGTACTGTCTTCTTCTGAGTACTGTCTGTCTGTGCAAGGCGGAGGTGATCAGTGA 1087  
 QY 1080 GCTCAGAGAGAGCTTCTGTCTTCTAAGTGCGCTGCAAGGCGGCACTCTTACTGTGAGTA 1139  
 Db 1088 GCTCAGAGAGAGCTTCTGTCTTCTAAGTGCGCTGCAAGGCGGCACTCTTACTGTGAGTA 1147  
 QY 1140 AGAGGTACCCACAGAAACCGGCTAGTGCAGAGAGTGTGCAAAAACAGACAAATGCAATG 1199  
 Db 1148 AGAGGTACCCACAGAAACCGGCTAGTGCAGAGAGTGTGCAAAAACAGACAAATGCAATG 1207  
 QY 1200 TGGAAATTTAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1259  
 Db 1208 TGGAAATTTAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1267  
 QY 1260 GATTTCACAGACTTAACCTTTTGTCTGTATATAAAGTACACCGTGTGTGTTTACATCTT 1319  
 Db 1268 GATTTCACAGACTTAACCTTTTGTCTGTATATAAAGTACACCGTGTGTGTTTACATCTT 1327  
 QY 1320 TGGGAAGCCAGGAAGGCTTGGAAAACGTATCACTTTCCAGATTTCTGGAATCTTCCA 1379  
 Db 1328 TGGGAAGCCAGGAAGGCTTGGAAAACGTATCACTTTCCAGATTTCTGGAATCTTCCA 1387  
 QY 1380 CTCTTTGCCAACAGCACTTGTCTGCGAACTCTTCTGGAATGCAATTCACATCAGATCCCC 1439  
 Db 1388 CTCTTTGCCAACAGCACTTGTCTGCGAACTCTTCTGGAATGCAATTCACATCAGATCCCC 1447  
 QY 1440 AACCGTGCAAGGTATCTGTCTTGTGCAAAAGAGTGTGATGTGAAATTCCTCTGTAG 1499  
 Db 1448 AACCGTGCAAGGTATCTGTCTTGTGCAAAAGAGTGTGATGTGAAATTCCTCTGTAG 1507  
 QY 1500 AATTAGCTTATACATTTACAGAAATAGCAGTTTCACTGCGCACTTTTATGTTGGTGAGAA 1559  
 Db 1508 AATTAGCTTATACATTTACAGAAATAGCAGTTTCACTGCGCACTTTTATGTTGGTGAGAA 1567  
 QY 1560 ATTTTAGTTAGTGTGTTGGATCGACCTCAGTTTCTGTGTTCTTTTATGTCGTGCT 1619  
 Db 1568 ATTTTAGTTAGTGTGTTGGATCGACCTCAGTTTCTGTGTTCTTTTATGTCGTGCT 1627  
 QY 1620 TTTCTATACATGATCATATGCCAAAACCTTTTGGAAAACGTTGTGAGATGATGTTGTT 1679  
 Db 1628 TTTCTATACATGATCATATGCCAAAACCTTTTGGAAAACGTTGTGAGATGATGTTGTT 1687  
 QY 1680 CTTTACCACCAAGAGATCAAGATACCTTGTATATAAAGCTGATGATATATTCAT 1739  
 Db 1688 CTTTACCACCAAGAGATCAAGATACCTTGTATATAAAGCTGATGATATATTCAT 1747  
 QY 1740 ACTGTGTGACACTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTAAC 1799

Db 1748 ACCGTGTGACCTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTAAC 1807  
 QY 1800 TGTGATCATATGTTTATGAGAAAAGCTGTTCCTCATGTCTTAACGAGACTGAATTCAAAAC 1859  
 Db 1808 TGTGATCATATGTTTATGAGAAAAGCTGTTCCTCATGTCTTAACGAGACTGAATTCAAAAC 1867  
 QY 1860 ATGTCAAGTGAATGATGATCTGTGCGATATGAGAGGATGCAAGTGCCTTCCCATTC 1919  
 Db 1868 ATGTCAAGTGAATGATGATCTGTGCGATATGAGAGGATGCAAGTGCCTTCCCATTC 1927  
 QY 1920 ATTCTGATGGAATGTTATACATGATGTTAACTGTGATATTTTCTGATGTATATG 1979  
 Db 1928 ATTCTGATGGAATGTTATACATGATGTTAACTGTGATATTTTCTGATGTATATG 1987  
 QY 1980 TATGCTGTGAATAGGATATATATTTTGGCTTAAACATACCGTAAATGTTGTGAT 2039  
 Db 1988 TATGCTGTGAATAGGATATATATTTTGGCTTAAACATACCGTAAATGTTGTGAT 2047  
 QY 2040 TTGAATACTTAAATGCGCAAGTAAATGCAATGCTTTGGAAATTTGGAAAGATGTTTAT 2099  
 Db 2048 TTGAATACTTAAATGCGCAAGTAAATGCAATGCTTTGGAAATTTGGAAAGATGTTTAT 2107  
 QY 2100 CTTGAGAGCAATATGTTTGCATTAATGCTTGAATGTTGATATCAAGAAATGAT 2159  
 Db 2108 CTTGAGAGCAATATGTTTGCATTAATGCTTGAATGTTGATATCAAGAAATGAT 2167  
 QY 2160 GAAGCTTCTCAAAACCTGTGTTACGTAAGTGTGAAGAGGAGCCGTTGGAGAGACCA 2219  
 Db 2168 GAAGCTTCTCAAAACCTGTGTTACGTAAGTGTGAAGAGGAGCCGTTGGAGAGACCA 2227  
 QY 2220 TTGCATGCTGTCCAGATGTTTCTGTTTAAAGCTTTTAACTGSAAGGCTAACCTCAA 2279  
 Db 2228 TTGCATGCTGTCCAGATGTTTCTGTTTAAAGCTTTTAACTGSAAGGCTAACCTCAA 2287  
 QY 2280 AATACCTTTTAACTGCACTTCTAATTAATGAGGCAAGATATCTCTTAC 2331  
 Db 2288 AATATTTTCTTAACTGCACTTCTAATTAATGAGGCAAGATATCTCTTAC 2339

RESULT 3  
 ID AAD30156 standard; DNA; 2212 BP.  
 AC AAD30156;  
 DT 17-MAY-2002 (first entry)  
 DE Human MCIP associated DNA #2.  
 XX  
 XX Human muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 25..618  
 FT /tag= a  
 FT /product= "Human MCIP associated protein #2"  
 PN MO200204491-A2.  
 PD 17-JAN-2002.  
 XX  
 XX 06-JUL-2001; 2001WO-US21662.  
 XX 07-JUL-2001; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXAS) UNIV TEXAS SYSTEM.  
 PA (WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX

PI Williams SR, Rothermel B;  
DR WPI; 2002-179698/23.  
DR P-PSDB; AAE18915.  
XX  
PT Screening for modulators of muscle calcineurin interacting protein  
PT (MCIP) binding, expression or phosphorylation, useful for treating  
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
PT calcineurin and a test compound -  
XX  
XX Example 1; Page 159-161; 174pp; English.  
XX  
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical  
XX complex with the catalytic subunit of calcineurin and increased levels  
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate  
XX transcription of certain target genes. The invention also relates to  
XX methods for identifying modulators of MCIP binding, expression or  
XX phosphorylation. Inhibitors or promoters of MCIP binding, expression or  
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
XX may be used for treating cardiac hypertrophy and heart failure.  
XX Antibodies to MCIP can be used in characterizing the MCIP content of  
XX healthy and diseased tissues and subsequently for determining the  
XX presence or absence of cardiomyopathy or as predictor of heart disease.  
XX The present invention is a human MCIP associated DNA.  
XX Note: This sequence has been described as murine MCIP splice variant in  
XX the specification, however the sequence seems to be a polynucleotide  
XX encoding a MCIP associated protein.  
XX  
SQ Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;  
  
Query Match 90.5%; Score 2109; DB 24; Length 2212;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 2142; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
  
QY 135 GAAAGCAAGATGCAATTTTAGAACTTTAACTACAGTTTGTAGCTCCCTGATTCCTGTG 194  
DB 16 GACTCGCAGATGGAGAGGTGGACCTCGAGGACTGCGCCAGCGCCACCATCGCCTGTAC 75  
  
QY 195 GCAACAGTATATCTTCAGGAAAGTGAACCAAGGCCAAATTTGAGTCCCTCTTTAGG 254  
DB 76 CTGACCCCGGCGGTTCGTGGACGGCGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGG 135  
  
QY 255 ACGTATGCAAGGACATCACCCTTCAGTATTTTAAAGCTTCAACAGAGTCAGAAATAAC 314  
DB 136 ACGTATGCAAGGACATCACCCTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGAAATAAC 195  
  
QY 315 TTGAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTCTG 374  
DB 196 TTGAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTCTG 255  
  
QY 375 GGAAGGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACCTGGCT 434  
DB 256 GGAAGGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACCTGGCT 315  
  
QY 435 CCGCCAAATCCAGCAAGCAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGG 494  
DB 316 CCGCCAAATCCAGCAAGCAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGG 375  
  
QY 495 AAACAAGTGAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCCAAG 554  
DB 376 AAACAAGTGAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCCAAG 435  
  
QY 555 CTGGGCGCCAGGGAAGATGAAATTTGACGAGCGAGCTGACACCACTCCAGCGGTGGT 614  
DB 436 CTGGGCGCCAGGGAAGATGAAATTTGACGAGCGAGCTGACACCACTCCAGCGGTGGT 495  
  
QY 615 GTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAGAGAGGAAATGGAAGATGAGG 674  
DB 496 GTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAGAGGAAATGGAAGATGAGG 555  
  
QY 675 AGACCTAAGCCAAATTTATCCAGACGAGGCGCGAGTACACGCCGATCCACCTCAGC 734  
DB 556 AGACCTAAGCCAAATTTATCCAGACGAGGCGCGAGTACACGCCGATCCACCTCAGC 615

QY 735 TGAAGTGGCAGCAGCAGAGCGCATTTCAAATCATACTACGCGGAGGAATCTTTTACTG 794  
DB 616 TGAAGTGGCAGCAGCAGAGCGCATTTCAAATCATACTACGCGGAGGAATCTTTTACTG 675  
  
QY 795 TCGAGGTGGCTGTGTACAGCTTCTTCGGAGGTGGCAGCGAGATCGGGTGGCAGAAATC 854  
DB 676 TCGAGGTGGCTGTGTACAGCTTCTTCGGAGGTGGCAGCGAGATCGGGTGGCAGAAATC 735  
  
QY 855 CAGTTTCATGTTGCTCTCAGAGAGAAATCAAGGCGGTGTCCCTTTGTTCTTAATGCTGCACAC 914  
DB 736 CAGTTTCATGTTGCTCTCAGAGAGAAATCAAGGCGGTGTCCCTTTGTTCTTAATGCTGCACAC 795  
  
QY 915 CAGTTACTGTTCATGCGACCCGGGATGACTTGGGCCAATCACTGAGTTGTGTGATCG 974  
DB 796 CAGTTACTGTTCATGCGACCCGGGATGACTTGGGCCAATCACTGAGTTGTGTGATCG 855  
  
QY 975 CACAAGGACATTTGGGACTGTCTTGAGAAAACAGATAATGATAGTGTGTTTGTACTTGTTC 1034  
DB 856 CACAAGGACATTTGGGACTGTCTTGAGAAAACAGATAATGATAGTGTGTTTGTACTTGTTC 915  
  
QY 1035 TTTTCTGTAGGTTCGTCTGTGCCAAGGCCAGGTGTGATCAGTGAGCTCAGGAGAGAGCT 1094  
DB 916 TTTTCTGTAGGTTCGTCTGTGCCAAGGCCAGGTGTGATCAGTGAGCTCAGGAGAGAGCT 975  
  
QY 1095 TCTGTGTTCTTAAGTGGCCTGCGAGGGGCCACTCTCTACTGGTAGGAGAGGTACCAACAG 1154  
DB 976 TCTGTGTTCTTAAGTGGCCTGCGAGGGGCCACTCTCTACTGGTAGGAGAGGTACCAACAG 1035  
  
QY 1155 AGCCGCTAGTGCAGAGAGGTTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTAGCT 1214  
DB 1036 AGCCGCTAGTGCAGAGAGGTTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTAGCT 1095  
  
QY 1215 TTCTCTTCTTCCCTCATGTTCTCATGTTTGTGTCATGTATATTAAGTAAAGCTAA 1274  
DB 1096 TTCTCTTCTTCCCTCATGTTCTCATGTTTGTGTCATGTATATTAAGTAAAGCTAA 1155  
  
QY 1275 CTTTGTGTTGTAATAAAGTTTACACGGTTGTTTGTGTTTACATCTTTTGGGAAGCCAGGAA 1334  
DB 1156 CTTTGTGTTGTAATAAAGTTTACACGGTTGTTTGTGTTTACATCTTTTGGGAAGCCAGGAA 1215  
  
QY 1335 GCGTTTGGAAAACGATCACCTTTCCAGATTTCCGAGTTCTCGACTCTTTGCAACAGCA 1394  
DB 1216 GCGTTTGGAAAACGATCACCTTTCCAGATTTCCGAGTTCTCGACTCTTTGCAACAGCA 1275  
  
QY 1395 CTTGCTTCGGGAACCTTCTCTGGAATGCAATTCATCAGATCCCAACCGTGCAACGTTG 1454  
DB 1276 CTTGCTTCGGGAACCTTCTCTGGAATGCAATTCATCAGATCCCAACCGTGCAACGTTG 1335  
  
QY 1455 AACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAATTCCTCTGTAGAAATTTAGTTTAGTG 1514  
DB 1336 AACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAATTCCTCTGTAGAAATTTAGTTTAGTG 1395  
  
QY 1515 APTCAGAGAAATGACAGTTTCACTGCCAACTTTAGTGGGTGAGAAAATTTAGTTTAGTG 1574  
DB 1396 APTCAGAGAAATGACAGTTTCACTGCCAACTTTAGTGGGTGAGAAAATTTAGTTTAGTG 1455  
  
QY 1575 TTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGTTCATACATGAAATC 1634  
DB 1456 TTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGTTCATACATGAAATC 1515  
  
QY 1635 ATAGCCAAAATCTTTTGGAAAACCTGTTGGTTGAGATAGTTGGTTCTTTTACCCACGAA 1694  
DB 1516 ATAGCCAAAATCTTTTGGAAAACCTGTTGGTTGAGATAGTTGGTTCTTTTACCCACGAA 1575  
  
QY 1695 GACATCAAGATACACTTGTGTAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1754  
DB 1576 GACATCAAGATACACTTGTGTAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1635  
  
QY 1755 GGGTGAAGAGTATGCGCAGTGGGAGACTTAAGATGTATTAACCTACCTGTGTAATCATATGTT 1814  
DB 1636 GGGTGAAGAGTATGCGCAGTGGGAGACTTAAGATGTATTAACCTACCTGTGTAATCATATGTT 1695

OY	1815	GTTGGAAAAGCGTTTCCCATGCTTAAACGAGACTTGAATTCAAAAGCATGCAAGTGATAG	1874
Db	1696	GTAGAAAAAGCGTTCGCCATGTCCTAACGAGACTTGAATTCAAAAGCATGCAAGTGATAG	1755
OY	1875	TAGATCTGTGGCGAATATGAGAGGATGCAAGTGCCTTTCCCATTCATTCTGATGGAAAT	1934
Db	1756	TAGATCTGTGGCGAATATGAGAGGATGCAAGTGCCTTTCCCATTCATTCTGATGGAAAT	1815
OY	1935	GTTATATCTAGGTTAACATTGTGAATTTTTTTCTAGTTGTAATGCTATGTCGTGAATA	1994
Db	1816	GTTATATCTAGGTTAACATTGTGAATTTTTTTCTAGTTGTAATGCTATGTCGTGAATA	1875
OY	1995	GGTATATATTTTGGCCCTTAAATACCCGTAACAAATGTTTGTCATTTTGGAAATCTTAATG	2054
Db	1876	GGTATATATATTTTGGCCCTTAAATACCCGTAACAAATGTTTGTCATTTTGGAAATCTTAATG	1935
OY	2055	CCAACTAACAAATGCATGCTTTGGAATTTGGAAGATGCTTTATTTCTTTGAGAACCAAT	2114
Db	1936	CCAACTAACAAATGCATGCTTTGGAATTTGGAAGATGCTTTATTTCTTTGAGAACCAAT	1995
OY	2115	ATGTTTGCAATTAATGCTTTGATTTGTTCAATTCAGAAATATGATGAACGTTCTCAACC	2174
Db	1996	ATGTTTGCAATTAATGCTTTGATTTGTTCAATTCAGAAATATGATGAACGTTCTCAACC	2055
OY	2175	CTGTTTACGGTACTTGTTAAGAGAGGAGCCGGTTTGGGAGAGACCAATGCAATGCTGTCCA	2234
Db	2056	CTGTTTACGGTACTTGTTAAGAGAGGAGCCGGTTTGGGAGAGACCAATGCAATGCTGTCCA	2115
OY	2235	ACTGTTTCTGTTAAGTGCCTTTTAAACTGAGAGGCTAACTCAAAATATCTTTTTTTAACT	2294
Db	2116	ACTGTTTCTGTTAAGTGCCTTTTAAACTGAGAGGCTAACTCAAAATATCTTTTTTTAACT	2175
OY	2295	TGCATTCTATTAATTAATTAATGGGCAACAGATGCTCTCTTAC	2331
Db	2176	TGCATTCTATTAATTAATTAATGGGCAACAGATGCTCTCTTAC	2212
RESULT 4			
ABN96843			
ID	ABN96843	standard; DNA; 2173 BP.	
AC	ABN96843;		
XX			
DT	13-AUG-2002	(first entry)	
XX			
DE			
XX			
KW	Gene #3341 used to diagnose liver cancer.		
KW	Gene; liver cancer; de; hepatocellular carcinoma; hepatotropic;		
KW	metastatic liver tumour; cyrobratic; expression profile; disease strate		
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.		
OS	Homo sapiens.		
XX			
PN	WO200229103-A2.		
XX			
PD	11-APR-2002.		
XX			
PF	02-OCT-2001; 2001WO-US30589.		
XX			
PR	02-OCT-2000; 2000US-237054P.		
XX			
PA	(GENE-) GENE LOGIC INC.		
XX			
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;		
XX	WPI, 2002-426119/45.		
DR			
XX			
PT	Diagnosing and detecting the progression of liver cancer,		
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,		
PT	involves detecting the level of expression of two or more genes in a		
XX	liver tissue sample -		
XX			
PS	Claim 1; SEQ ID NO 3341; 298bp; English.		

CC	XX	The invention relates to a novel method for diagnosing and detecting the			
CC	XX	progression of liver cancer, hepatocellular carcinoma or metastatic liver			
CC	XX	tumour in a patient, and differentiating metastatic liver cancer from			
CC	XX	hepatocellular carcinoma in a patient, involving detecting the level of			
CC	XX	expression of two or more genes represented in ABN93503-ABN97455 in a			
CC	XX	tissue sample. The method of the invention has hepatotropic, and			
CC	XX	cytostatic activity. The method is useful for diagnosing and detecting			
CC	XX	the progression of liver cancer, hepatocellular carcinoma and metastatic			
CC	XX	liver carcinoma in a patient. The method is useful for identifying			
CC	XX	expression profiles which serve as useful diagnostic markers as well as			
CC	XX	markers that can be used to monitor disease states, disease progression,			
CC	XX	drug toxicity, drug efficacy and drug metabolism.			
CC	XX	Note: The sequence data for this patent did not form part of the printed			
CC	XX	specification, but was obtained in electronic format directly from WIPO			
CC	XX	at <a href="http://wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a> .			
CC	XX				
CC	XX	Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other:			
CC	XX				
CC	XX	Query Match 90.1%; Score 2101; DB 24; Length 2173;			
CC	XX	Best Local Similarity 100.0%; Pred. NO. 0;			
CC	XX	Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0			

Query Match	Similarity	100.0%	Score 2101	DB 24	Length 2173
Best Local	Similarity	100.0%	Pred. No. 0		
Matches 2101	Conservative	0	Mismatches	0	Indels
					Gaps
Qy 231	GCCTAAATTTGAGTCCCTCTTTAGACGATGACAAAGACATCACCTTTCAGTATTTTAAAG	230			
Db 58	GCCTAAATTTGAGTCCCTCTTTAGACGATGACAAAGACATCACCTTTCAGTATTTTAAAG	117			
Qy 291	AGCTTCAAAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGCATGCCAGGCTC	350			
Db 118	AGCTTCAAAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGCATGCCAGGCTC	177			
Qy 351	CAGCGCATTAAGCTGAGTTCGGGAAAGAAATGAATATATTTTGTCCAGACCTTA	410			
Db 178	CAGCGCATTAAGCTGAGTTCGGGAAAGAAATGAATATATTTTGTCCAGACCTTA	237			
Qy 411	CACATAGGAAGCTCACACTCGGCTCCGCCAAATCCAGACAGCACTTTCGATCTCCCT	470			
Db 238	CACATAGGAAGCTCACACTCGGCTCCGCCAAATCCAGACAGCACTTTCGATCTCCCT	297			
Qy 471	CCCGCTCTCCGCCAGTGGAGTGAACAAATGGAAGATGCAACCCACGTATTAATCTAT	530			
Db 298	CCCGCTCTCCGCCAGTGGAGTGAACAAATGGAAGATGCAACCCACGTATTAATCTAT	357			
Qy 531	GATCTCTTATATGCTCATCTCCAAAGTGGGGCCAGGGGAAATATGAATTCACACGACG	590			
Db 358	GATCTCTTATATGCTCATCTCCAAAGTGGGGCCAGGGGAAATATGAATTCACACGACG	417			
Qy 591	ACTGACACCACTCCACAGCGTGTGTGTCATGTATGTGACAGTGCATCAAGACAGGAGAA	650			
Db 418	ACTGACACCACTCCACAGCGTGTGTGTCATGTATGTGACAGTGCATCAAGACAGGAGAA	477			
Qy 651	GAAAGAGAAATGGAAGATGAGAGACCTTAACCAAAAATATTCAGACCGAGAGGCG	710			
Db 478	GAAAGAGAAATGGAAGATGAGAGACCTTAACCAAAAATATTCAGACCGAGAGGCG	537			
Qy 711	GAGTACACGCGGATCCACTCACTGACTGAATCTGGCACGCGACAGAGAGCATTTCCAAATAT	770			
Db 538	GAGTACACGCGGATCCACTCACTGACTGAATCTGGCACGCGACAGAGAGCATTTCCAAATAT	597			
Qy 771	ACTCACGGGAGGAATCTTTTACGTGGAGGTGCTGTCACGACTTCTCGAGGTGGCA	830			
Db 598	ACTCACGGGAGGAATCTTTTACGTGGAGGTGCTGTCACGACTTCTCGAGGTGGCA	657			
Qy 831	GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGTG	890			
Db 658	GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGTG	717			
Qy 891	TCCCTTGTCTTAATGCTGCACACAGTTACTGTTCAATGACCCGGGATATGACTTTGGGC	950			
Db 718	TCCCTTGTCTTAATGCTGCACACAGTTACTGTTCAATGACCCGGGATATGACTTTGGGC	777			
Qy 951	CAATCACTGAGTTTGTGTGATGCCACAAAGACATTTTGGGACCTGCTTTAGAAAACAGAT	1010			



Db 778\* CAATCACTGAGTTTGTGGTATCGCACAGGACATTTGGGACTGTCTTGAGAAAACGAT 837  
QY 1011 AATGATAGTGTGTTTGTACTTGTCTTTCTGGTAGGTTCTGTCTGTGCAAGGCGAGTT 1070  
Db 838 AATGATAGTGTGTTTGTACTTGTCTTTCTGGTAGGTTCTGTCTGTGCAAGGCGAGTT 897  
QY 1071 GATCAGTGTAGCTCAGGAGAGCTTCTGTTTCTAAGTGGCTGTGAGGGCCACTCTCTA 1130  
Db 898 GATCAGTGTAGCTCAGGAGAGCTTCTGTTTCTAAGTGGCTGTGAGGGCCACTCTCTA 957  
QY 1131 CTGTAGGAAGAGGTACCAAGGAGCGCCCTAGTGCAGAGAGTTGTGAAAACAGGAGC 1190  
Db 958 CTGTAGGAAGAGGTACCAAGGAGCGCCCTAGTGCAGAGAGTTGTGAAAACAGGAGC 1017  
QY 1191 AATGCAATGTGCAATTTAGGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCAATG 1250  
Db 1018 AATGCAATGTGCAATTTAGGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCAATG 1077  
QY 1251 TATATTACTGATTTACAGAGCTAACCTTTGTTGCTATATATAAGTTTACACCGTTGTTGTTT 1137  
Db 1078 TATATTACTGATTTACAGAGCTAACCTTTGTTGCTATATATAAGTTTACACCGTTGTTGTTT 1137  
QY 1311 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTCG 1370  
Db 1138 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTCG 1197  
QY 1371 GATTCTGACTCTTTGCAAGCACTTCTGTCGGAAGTCTTCTCTGGAATGCATTCACCTC 1430  
Db 1198 GATTCTGACTCTTTGCAAGCACTTCTGTCGGAAGTCTTCTCTGGAATGCATTCACCTC 1257  
QY 1431 AGCATCCCAACCGTCAACGTTGAACTTGTGCTTTTGCAGAAAGTGTGATCTGAAAT 1490  
Db 1258 AGCATCCCAACCGTCAACGTTGAACTTGTGCTTTTGCAGAAAGTGTGATCTGAAAT 1317  
QY 1491 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCACTTCTGCACTGCACTTTAGT 1550  
Db 1318 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCACTTCTGCACTGCACTTTAGT 1377  
QY 1551 GGGTGAGAAATTTAGTATAGTGTGTTGGGATCGGACCTCAGTTCTGTGTTTCTTTTA 1610  
Db 1378 GGGTGAGAAATTTAGTATAGTGTGTTGGGATCGGACCTCAGTTCTGTGTTTCTTTTA 1437  
QY 1611 TGTGTTGTTTCTTATACATGAATCATAGCCAAAGCTTTTGTGGAACCTGTTGTTGAGA 1670  
Db 1438 TGTGTTGTTTCTTATACATGAATCATAGCCAAAGCTTTTGTGGAACCTGTTGTTGAGA 1497  
QY 1671 TAGTTGTTCTTTTACCCCAAGAGACATCAAGATACACTTGTAAATAAAGCTGATAGCA 1730  
Db 1498 TAGTTGTTCTTTTACCCCAAGAGACATCAAGATACACTTGTAAATAAAGCTGATAGCA 1557  
QY 1731 TATATTACATCTGTTGTACCTTGGGTGAAAGATATGCGAGTGGGAGACTAAGATGTAT 1790  
Db 1558 TATATTACATCTGTTGTACCTTGGGTGAAAGATATGCGAGTGGGAGACTAAGATGTAT 1617  
QY 1791 TAACTCTGTTGAAATCATATGTTGTAGAAAAGCTGTTCCCATGCTCAACAGGACTTGA 1850  
Db 1618 TAACTCTGTTGAAATCATATGTTGTAGAAAAGCTGTTCCCATGCTCAACAGGACTTGA 1677  
QY 1851 ATTCAAGCATGTCAAGTGTAGTATCTGTGGCGATATGAGAGGGATGCAAGTGCCTT 1910  
Db 1678 ATTCAAGCATGTCAAGTGTAGTATCTGTGGCGATATGAGAGGGATGCAAGTGCCTT 1737  
QY 1911 TCCCATTCATCTCTGATGGAATTTGTTACTAGGTTTAACTTTGTTTCTAGT 1970  
Db 1738 TCCCATTCATCTCTGATGGAATTTGTTACTAGGTTTAACTTTGTTTCTAGT 1797  
QY 1971 TGTATGTGTATGTCTGGTAAATAGGTATATATTTTGGCCCTTACAATACCGTAAACAATG 2030  
Db 1798 TGTATGTGTATGTCTGGTAAATAGGTATATATTTTGGCCCTTACAATACCGTAAACAATG 1857  
QY 2031 TTGTGCTATTTGAAATCTTATGCAAGTAAACAATGCAATGCAATGCTTTTGGAAATTTGGAAGT 2090  
Db 1858 TTGTGCTATTTGAAATCTTATGCAAGTAAACAATGCAATGCAATGCTTTTGGAAATTTGGAAGT 1917

QY 2091 GGTTTTATTTCTTTGAGAAGCAATATGTTTGCATTAAATGCTTTGATTGTTTCATATCAAG 2150  
Db 1918 GGTTTTATTTCTTTGAGAAGCAATATGTTTGCATTAAATGCTTTGATTGTTTCATATCAAG 1977  
QY 2151 AAATTGATTGAACGTTTCTCAAAACCCCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGG 2210  
Db 1978 AAATTGATTGAACGTTTCTCAAAACCCCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGG 2037  
QY 2211 GAGAGACCAATGCAATGCTGTCCAAAGTGTTCCTTGTTAAGTGTCTTTAAACTGGAGAGC 2270  
Db 2038 GAGAGACCAATGCAATGCTGTCCAAAGTGTTCCTTGTTAAGTGTCTTTAAACTGGAGAGC 2097  
QY 2271 TAACTCTCAAAATACTTTTAACTGCAATCTATATAAATGGGCAAGTATGCTTCTTAA 2330  
Db 2098 TAACTCTCAAAATACTTTTAACTGCAATCTATATAAATGGGCAAGTATGCTTCTTAA 2157  
QY 2331 C 2331  
Db 2158 C 2158

RESULT 5  
AAFI8328  
ID AAFI8328 standard; DNA; 2358 BP.  
XX AAFI8328;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
XX Lung cancer associated polynucleotide sequence SEQ ID 347.  
XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
KW cardioactive; immunomodulatory; muscular active; vulnerary;  
KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KW proliferative disorder; wound healing; infectious disease; ds.  
XX Homo sapiens.  
XX  
XX WO200055180-A2.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05918.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
XX  
XX Ruben SM;  
XX  
XX WPI; 2000-587514/55.  
XX P-PSDB; AAB58452.  
XX  
XX Lung cancer associated gene sequences, referred to as lung cancer  
XX antigens, useful for treatment, prevention, and diagnosis of disorders  
XX such as lung cancer -  
XX  
XX Claim 1; Page 805-806; 1425pp; English.  
XX  
XX Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer  
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer  
XX associated proteins and polynucleotide sequences, their agonists, and  
XX antagonists may have neuroprotective; cytostatic; cardioactive;  
XX immunomodulatory; muscular active general; vulnerary; gastrointestinal  
XX general; nephrotropic; antiinfective; gynecological; or antibacterial  
XX activity. The invention also includes antibodies specific for the  
XX protein or polynucleotide sequences. The lung cancer associated  
XX polynucleotide sequences may be used for detection of lung cancer,  
XX chromosome identification, as chromosome markers, and for numerous other  
XX diagnostic or research purposes. The proteins may be used to treat

disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are used in the course of the invention for the identification and characterization of the polynucleotide and protein sequences.

Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other:

Query Match 89.8%; Score 2094.2; DB 21; Length 2358;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 2137; Conservative 3; Mismatches 56; Indels 1; Gaps 1;

135 GAAAGCAAGATGATTTTAACTTAACCTTAAGCTTCCCTGATTTCCGCTGTG 194  
Db GACTGGAGATGAGAGAGTGGACCTGACAGACCTGCCAGGCCACCTCCCTGTAC 164  
195 GCAACAGATATCTCTTCAAGCGAAGTGAACCCAGGCCAAATTTGAGTCCCTTTAGG 254  
165 CTGGACCCCGCGCTGTTCTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTTTAGG 224  
255 ACCTATGACAGAGATCAACCTTTCAGTATTTTAAAGCTTCAACAGATCAAGATAAC 314  
225 ACCTATGACAGAGATCAACCTTTCAGTATTTTAAAGCTTCAACAGATCAAGATAAC 284  
315 TTGACGACCCCTTCTCCGACAGATGCCAGCTCCAGTGTGATAGACTGAGTTCTG 374  
285 TTGACGACCCCTTCTCCGACAGATGCCAGCTCCAGTGTGATAGACTGAGTTCTG 344  
375 GGAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGAGAGCTCACACTGCT 434  
345 GGAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGAGAGCTCACACTGCT 404  
435 CCGCAATTCAGACAGACGATTTCTGATCTCCCTCCGCTCTCCGACAGTGGATGG 494  
405 CCG-CAATTCAGACAGACGATTTCTGATCTCCCTCCGCTCTCCGACAGTGGATGG 463  
495 AACAAGTGAAGATGAGACCCAGCTATTAAGTATCTTATATGCACTCCCAAG 554  
464 AACAAGTGAAGATGAGACCCAGCTATTAAGTATCTTATATGCACTCCCAAG 523  
555 CTGGGCGCAGGGGAAAATGATGAATTCGACGACGCACTGACCACTCCAGCGTGTG 614  
524 CTGGGCGCAGGGGAAAATGATGAATTCGACGACGCACTGACCACTCCAGCGTGTG 583  
615 GTCCATATATGTAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674  
584 GTCCATATATGTAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643  
675 AGACTTAAGCCAAATAATTCAGACAGAGAGCGGAGTACAGCCGATCCAGCTCAGC 734  
644 AGACTTAAGCCAAATAATTCAGACAGAGAGCGGAGTACAGCCGATCCAGCTCAGC 703  
735 TGAATCTGACGACGACGACGACGATTCGAATCACTACGCGGAGAAATCTTTTCTG 794  
704 TGAATCTGACGACGACGACGACGATTCGAATCACTACGCGGAGAAATCTTTTCTG 763  
795 TGGAGTGTGCTGTCAAGACTTTCTCGAGAGTGGACGCGAGATCGGGGTGGAGAAATC 854  
764 TGGAGTGTGCTGTCAAGACTTTCTCGAGAGTGGACGCGAGATCGGGGTGGAGAAATC 823  
855 CCAAGTTCATGTTGCTCAGAAAGAGATCAAGCGTGTCCCTGTTCTAATGCTGCAAC 914  
824 CCAAGTTCATGTTGCTCAGAAAGAGATCAAGCGTGTCCCTGTTCTAATGCTGCAAC 883  
915 CAGTTCATGTTGCTCAGACCCCGGAGATGATTTGGCCAAATCACTGAGATTTGTGTGATCG 974  
884 CAGTTCATGTTGCTCAGACCCCGGAGATGATTTGGCCCAATCACTGAGATTTGTGTGATCG 943  
975 CACAAGGACATTTGGGACATGCTTGAAGAAAACAGATATGATAGTGTGTTGACTGTTTC 1034  
944 CACAAGGACATTTGGGACATGCTTGAAGAAAACAGATATGATAGTGTGTTGACTGTTTC 1003

1035 TTTCTGTGATGTTCTGTCTGTGTCACAGGCGAGTTGATCAGTACCTCAGAGAGAGCT 1094  
1004 TTTCTGTGATGTTCTGTCTGTGTCACAGGCGAGTTGATCAGTACCTCAGAGAGAGCT 1063  
1095 TCTGTGTTCTAAGTGGCTGACAGGGGCACTCTTCTGTGTAGAGAGAGTACCAAGAG 1154  
1064 TCTGTGTTCTAAGTGGCTGACAGGGGCACTCTTCTGTGTAGAGAGAGTACCAAGAG 1123  
1155 AGCGGCTAGTGCAGAGAGTGTGAAAACAGACAGATGCAATGTGAAATTTAGCGT 1214  
1124 AGCGGCTAGTGCAGAGAGTGTGAAAACAGACAGATGCAATGTGAAATTTAGCGT 1183  
1215 TTCTTTCTTCCCTCATGTTCTGATGTTGTGATGATATTAAGTATTTAACAAGCTAA 1274  
1184 TTCTTTCTTCCCTCATGTTCTGATGTTGTGATGATATTAAGTATTTAACAAGCTAA 1243  
1275 CTTTGTTCGATTTAAGTACACCGTGTGTTTAACTCTTTGGGAGGCGAGAGAA 1334  
1244 CTTTGTTCGATTTAAGTACACCGTGTGTTTAACTCTTTGGGAGGCGAGAGAA 1303  
1335 GCGTTTGGAAAACGATACACTTCCAGATTCCTCGAATTCGACTCTTTGACAGCA 1394  
1304 GCGTTTGGAAAACGATACACTTCCAGATTCCTCGAATTCGACTCTTTGACAGCA 1363  
1395 CTGCTTGGGAACCTTCTCTGAGATGCACTACGATCCCAACGCTGCAAGCTGT 1454  
1364 CTGCTTGGGAACCTTCTCTGAGATGCACTACGATCCCAACGCTGCAAGCTGT 1423  
1455 AACTTGTGCTTTGCAAAAAGAGTGAATCGAATTCCTCTGTGAATTTAGTTATACA 1514  
1424 AACTTGTGCTTTGCAAAAAGAGTGAATCGAATTCCTCTGTGAATTTAGTTATACA 1483  
1515 ATTGAGAAATGAGAGTTCATCTGCAACTTTTATGAGTGAAGAAATTTAGTTAGTG 1574  
1484 ATTGAGAAATGAGAGTTCATCTGCAACTTTTATGAGTGAAGAAATTTAGTTAGTG 1543  
1575 TTTGGATTCGACACTCAGTTCTGTGTTCTTTATGAGTGGTTCCTATACATGATC 1634  
1544 TTTGGATTCGACACTCAGTTCTGTGTTCTTTATGAGTGGTTCCTATACATGATC 1603  
1635 ATAGCCAAAACCTTTTGGAAAACGTTGTTGAGATAGTGTGTTTCTTTTACCCCAAGAA 1694  
1604 ATAGCCAAAACCTTTTGGAAAACGTTGTTGAGATAGTGTGTTTCTTTTACCCCAAGAA 1663  
1695 GACATCAAGATACACTTGTAAATTAAGCTGATAGCATATTAATCACTGTTGTACCTT 1754  
1664 GACATCAAGATACACTTGTAAATTAAGCTGATAGCATATTAATCACTGTTGTACCTT 1723  
1755 GGGTGAAGAAATGAGCACTGGGAGCTAAGATGATTAACCTACCTGAGATCATGTT 1814  
1724 GGGTGAAGAAATGAGCACTGGGAGCTAAGATGATTAACCTACCTGAGATCATGTT 1783  
1815 GTAGGAAAAGCTGTTCCCATGTTCTAAGCAGACTTGAAATTAAGCATGCAAGTGATG 1874  
1784 GTAGGAAAAGCTGTTCCCATGTTCTAAGCAGACTTGAAATTAAGCATGCAAGTGATG 1843  
1875 TAGATCTGTGGGATATGAGAGAGATGCAAGTCCCTTTCCCATTTCAATTCCTGATGGAAT 1934  
1844 TAGATCTGTGGGATATGAGAGAGATGCAAGTCCCTTTCCCATTTCAATTCCTGATGGAAT 1903  
1935 GTTATCTAGGTTAATCATTTGATTTTCTAGTGTAAATGATGATGCTGTAATA 1994  
1904 GTTATCTAGGTTAATCATTTGATTTTCTAGTGTAAATGATGATGCTGTAATA 1963  
1995 GGTATTATATTTTGGCTTACAAATCCGTAACAGTGTGTCATTTGAAATATCTTAATG 2054  
1964 GGTATTATATTTTGGCTTACAAATCCGTAACAGTGTGTCATTTGAAATATCTTAATG 2023  
2055 CCAAGTAAATGATGCTCTTGGAAATTTGGAAGATGTTTATCTTGGAGAACCAAT 2114  
2024 CCAAGTAAATGATGCTCTTGGAAATTTGGAAGATGTTTATCTTGGAGAACCAAT 2083



QY 2115 ATGTTTGCATTAATGCTTTGATTTGTTTCAATATCAAGAAATGATTAAGACCTTCTCAAAACC 2174  
 Db |||||||  
 QY 2084 AGTTTTCATTAATGCTTTGATTTGTTTCAATATCAAGAAATGATTAAGACCTTCTCAAAACC 2143  
 Db |||||||  
 QY 2175 CTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCAATTCGATCGCTGTCCA 2234  
 Db |||||||  
 QY 2144 CTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCAATTCGATCGCTGTCCA 2203  
 Db |||||||  
 QY 2235 ACTGTTTCTTGTAGTGTCTTTTAACTGGAGAGGCTAACCTCAAAATACCTTTTAAAC 2294  
 Db |||||||  
 QY 2204 AGTGTTCCTTGTAGTGTCTTTTAACTGGAGAGGCTAACCTCAAAATATTTTAAAC 2263  
 Db |||||||  
 QY 2295 TGCATTCCTATAATAATGGGACACGATGCTCCCTTAC 2331  
 Db |||||||  
 QY 2264 TGCATTCCTATAATAATGGGACACGATGCTCCCTTAC 2300  
 Db |||||||

RESULT 6  
 AAX01282  
 ID AAX01282 standard; cDNA to mRNA; 2174 BP.  
 XX  
 AC AAX01282;  
 XX  
 DT 09-APR-1999 (first entry)  
 XX  
 DE Human DSCR1 coding sequence.  
 XX  
 KW DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;  
 KW Central Nervous System development; mental retardation; heart defect; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 49..564  
 FT /\*tag= a  
 FT polyA\_signal 1541..1546  
 FT /\*tag= b  
 FT polyA\_signal 2132..2137  
 FT /\*tag= c  
 XX  
 PN US5869318-A.  
 XX  
 XX 09-FEB-1999.  
 PD  
 XX 07-JUN-1996; 96US-0665040.  
 XX  
 XX 07-JUN-1995; 95ES-0001140.  
 XX  
 PA (PALL/) PALLEJA X E.  
 XX  
 PI Puentes JJ, Palleja XE, Pritchard M;  
 XX  
 DR WPI; 1999-152781/13.  
 DR P-PSDB; AAW73898.  
 XX  
 PT DNA encoding foetal brain proteins - believed to be associated with  
 PT Down's syndrome  
 XX  
 XX Claim 1; Column 15-18; 19pp; English.  
 CC  
 CC This sequence is the Down's Syndrome critical region 1 (DSCR1) gene  
 CC of the invention. The DSCR1 gene was found to be located in the  
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient  
 CC expression of DSCR1 mRNA in the brains of young rats, compared to  
 CC expression levels in the brains of adult rats, suggests an important role  
 CC for DSCR1 during the development of the Central Nervous System (CNS), and  
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities  
 CC of mental retardation and/or heart defects as found in Down's syndrome  
 CC patients.  
 XX  
 SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 528 T; 3 other;  
 Query Match 88.3%; Score 2059.2; DB 20; Length 2174;

Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 2081; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 231 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGACATCACCTTTTCAGTATTTTAAAG 290  
 Db |||||||  
 QY 291 AGCTTCAGGAGTCAAGATAAATTTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 350  
 Db |||||||  
 QY 118 AGCTTCAGGAGTCAAGATAAATTTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177  
 Db |||||||  
 QY 351 CAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTATATTTTCTCAGACCTTA 410  
 Db |||||||  
 QY 178 CAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTATATTTTCTCAGACCTTA 237  
 Db |||||||  
 QY 411 CACATAGGAGCTCACACCTGGCTCCGCAATTCAGACAAGCAGTTTCTGATCTCCCT 470  
 Db |||||||  
 QY 238 CACATAGGAGCTCACACCTGGCTCCGCAATTCAGACAAGCAGTTTCTGATCTCCCT 297  
 Db |||||||  
 QY 471 CCGGCTCTCCGCGAGTGGGATGGAACAAGTGGAGATGGACCCAGTCATAAATAT 530  
 Db |||||||  
 QY 298 CCGGCTCTCCGCGAGTGGGATGGAACAAGTGGAGATGGACCCAGTCATAAATAT 357  
 Db |||||||  
 QY 531 GATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTCAGACGCG 590  
 Db |||||||  
 QY 358 GATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTCAGACGCG 417  
 Db |||||||  
 QY 591 ACTGACACACTCCGAGCGTGGTGCATGATGTGAGAGTGTATCAAGAGAAGAGGAA 650  
 Db |||||||  
 QY 418 ACTGACACACTCCGAGCGTGGTGCATGATGTGAGAGTGTATCAAGAGAAGAGGAA 477  
 Db |||||||  
 QY 651 GAAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAATATTCAGACGAGGCGG 710  
 Db |||||||  
 QY 478 GBAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAATATTCAGACGAGGCGG 537  
 Db |||||||  
 QY 711 GAGTACACGCGGATCCACTCAGCTGAACTGGCACGCGGACGAGCAGCATTCCAAATCAT 770  
 Db |||||||  
 QY 538 GAGTACACGCGGATCCACTCAGCTGAACTGGCACGCGGACGAGCAGCATTCCAAATCAT 597  
 Db |||||||  
 QY 771 ACTCAGCGGAGGAAATCTTTTACTGTGAGAGTGGTGTGTCAGCATTTCTGGAGGTGGCA 830  
 Db |||||||  
 QY 598 ACTCAGCGGAGGAAATCTTTTACTGTGAGAGTGGTGTGTCAGCATTTCTGGAGGTGGCA 657  
 Db |||||||  
 QY 831 GCCAGATCGGGGTGGCAGAAATCCCAAGTTCATGCTCAGAGAGAAATCAAGGCCGTG 890  
 Db |||||||  
 QY 658 GCCAGATCGGGGTGGCAGAAATCCCAAGTTCATGCTCAGAGAGAAATCAAGGCCGTG 717  
 Db |||||||  
 QY 891 TCCCTTGTCTTAATGTGCAACCACTTACTGTTTCATGGCACCCCGGGAATGACTTGGGC 950  
 Db |||||||  
 QY 718 TCCCTTGTCTTAATGTGCAACCACTTACTGTTTCATGGCACCCCGGGAATGACTTGGGC 777  
 Db |||||||  
 QY 951 CAATCACTGAGTTTGTGGTGTGTCAGCAAGACATTTGGGACTGTCTTGAGAAAACAGAT 1010  
 Db |||||||  
 QY 778 CAATCACTGAGTTTGTGGTGTGTCAGCAAGACATTTGGGACTGTCTTGAGAAAACAGAT 837  
 Db |||||||  
 QY 1011 AATGATAGTGTTTGTACTTTGTTCTTTCTGGTAGGTTCTGCTGTGCCAAGGCGAGTT 1070  
 Db |||||||  
 QY 838 AATGATAGTGTTTGTACTTTGTTCTTTCTGGTAGGTTCTGCTGTGCCAAGGCGAGTT 897  
 Db |||||||  
 QY 1071 GATCAGTGAGCTCAGGAGAGGCTTCTGTTCTTAAGTGGCTTCAGGGGCCACTCTCTA 1130  
 Db |||||||  
 QY 898 GATCAGTGAGCTCAGGAGAGGCTTCTGTTCTTAAGTGGCTTCAGGGGCCACTCTCTA 957  
 Db |||||||  
 QY 1131 CTGGTAGGAGAGGTACCAAGAGCCGCTAGTCGAGAGGTTGTGAAAACAGCAGC 1190  
 Db |||||||  
 QY 958 CTGGTAGGAGAGGTACCAAGAGCCGCTAGTCGAGAGGTTGTGAAAACAGCAGC 1017  
 Db |||||||  
 QY 1191 AATGCAATGGAATTTAGCGTTTCTTTCTTCCTCATGTTCTCATGTTTGTGTCATG 1250  
 Db |||||||  
 QY 1018 AATGCAATGGAATTTAGCGTTTCTTTCTTCCTCATGTTCTCATGTTTGTGTCATG 1077  
 Db |||||||  
 QY 1251 TATATTACTGATTTTACAAGACTTAACCTTTGTCGTATATAAAGTTTACCGTTTGTGTTT 1310  
 Db |||||||

Db 1078 TATATTACTGATTTTCAAGACTAACCTTGTCTGTATATAAGTTACACGGTGTGTTT 1137  
 Oy 1311 TACATCTTTGGGAAGCCAGAAAGCGTTTGGAAAAGTATACCTTTCCAGATTCG 1370  
 Db 1138 TACATCTTTGGGAAGCCAGAAAGCGTTTGGAAAAGTATACCTTTCCAGATTCG 1197  
 Oy 1371 GATTCTCGACTCTTTGGCAACAGCACTTGTGGGAACTCTTCTGGAATGATTCATCTC 1430  
 Db 1198 GATTCTCGACTCTTTGGCAACAGCACTTGTGGGAACTCTTCTGGAATGATTCATCTC 1257  
 Oy 1431 AGCATCCCAACCGTGCAACAGTGTACCTTGRCCTTTGCAAAAAGATGATCTGAAAAT 1490  
 Db 1258 AGCATCCCAACCGTGCAACAGTGTACCTTGRCCTTTGCAAAAAGATGATCTGAAAAT 1317  
 Oy 1491 CCTGTGTAAGATTAGCTTATACATTCAGAGATAGCACTTTCACCTGCACTTTTGT 1550  
 Db 1318 CCTGTGTAAGATTAGCTTATACATTCAGAGATAGCACTTTCACCTGCACTTTTGT 1377  
 Oy 1551 GGGTGGAAATTTTATAGTGTGGGATCGGACCTCAGTTCTGTGTCTTTTCTTTA 1610  
 Db 1378 GGGTGGAAATTTTATAGTGTGGGATCGGACCTCAGTTCTGTGTCTTTTCTTTA 1437  
 Oy 1611 TGTGTGTGTCTTATACATTCAGATTCAGCAAAACCTTTTGGAAAAGTGTGTGAGA 1670  
 Db 1438 TGTGTGTGTCTTATACATTCAGATTCAGCAAAACCTTTTGGAAAAGTGTGTGAGA 1497  
 Oy 1671 TACTGTGTTCTTTTACCCCGCAAGACATCAGATACACTGTGTAATTAAGCGTATGAGA 1730  
 Db 1498 TACTGTGTTCTTTTACCCCGCAAGACATCAGATACACTGTGTAATTAAGCGTATGAGA 1557  
 Oy 1731 TATATTCATCTGTGTGTACACTTGGGTGAAAAGTATGGCAGTGGAGACTAAGATGAT 1790  
 Db 1558 TATATTCATCTGTGTGTACACTTGGGTGAAAAGTATGGCAGTGGAGACTAAGATGAT 1617  
 Oy 1791 TAACTACCTGTGAATCATATGTGTGAGAAAAGCTGTCTCCATGTCTTAACGAGATTA 1850  
 Db 1618 TAACTACCTGTGAATCATATGTGTGAGAAAAGCTGTCTCCATGTCTTAACGAGATTA 1677  
 Oy 1851 ATTCAAGACATGCAAGTGAATGTAGATCTGTGGGCAATGAGAGGAGTACGTCCTT 1910  
 Db 1678 ATTCAAGACATGCAAGTGAATGTAGATCTGTGGGCAATGAGAGGAGTACGTCCTT 1737  
 Oy 1911 TCCCATTCATCTGTGTGTATCTAGTATACATTTGTAATTTTTTCTAGT 1970  
 Db 1738 TCCCATTCATCTGTGTGTATCTAGTATACATTTGTAATTTTTTCTAGT 1797  
 Oy 1971 TGTATGTGTATGTGTGTGAATGTATTTATTTTGGCTTACATACCGTAACAAG 2030  
 Db 1798 TGTATGTGTATGTGTGTGAATGTATTTATTTTGGCTTACATACCGTAACAAG 1857  
 Oy 2031 TTTTGCATTTGGAATCTTAAATGCAATGCAATGCAATGCAATTTGGAAATTTGGAGAT 2090  
 Db 1858 TTTTGCATTTGGAATCTTAAATGCAATGCAATGCAATGCAATTTGGAAATTTGGAGAT 1917  
 Oy 2091 GGTTTTATCTTTGGAAGCAATATGTGTGATTAATGCTTTGATTTGTCATATCAAG 2150  
 Db 1918 GGTTTTATCTTTGGAAGCAATATGTGTGATTAATGCTTTGATTTGTCATATCAAG 1977  
 Oy 2151 AAATGATTTGAACGTTCTCAAAACCTGTGTAAGGATCTGTAAGAGGAGCCGTTTG 2210  
 Db 1978 AAATGATTTGAACGTTCTCAAAACCTGTGTAAGGATCTGTAAGAGGAGCCGTTTG 2037  
 Oy 2211 GAGAGACCATTCATCGCT-GTCCAAAGTCTTCTGTGTAAGTCTTTTAACTGAGAGG 2269  
 Db 2038 GAGAGACCATTCATCGCTGTCNCAAGTCTTCTGTGTAAGTCTTTTAACTGAGAGG 2097  
 Oy 2270 CTAACTCAAAATACCTTTTAACTGATCTTAAATTAATGAGCAAGATATCTCTT 2329  
 Db 2098 CTAACTCAAAATACCTTTTAACTGATCTTAAATTAATGAGCAAGATATCTCTT 2157  
 Oy 2330 AC 2331  
 Db 2158 AC 2159

RESULT 7  
 ID ABA42132 standard; DNM; 486 BP.  
 XX  
 AC ABA42132;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human breast cell single exon nucleic acid probe #827.  
 XX  
 KW Human; microarray; single exon probe; gene expression; breast;  
 KW disease; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157271-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001MO-US00662;  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0508408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI, 2001-496933/54.  
 XX  
 PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes -  
 XX  
 PS Claim 1; SEQ ID NO 827; 327bp + sequence listing; English.  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene  
 CC expression analysis is useful for assessing the toxicity of chemical  
 CC agents on cells. The microarray of this invention presents a far greater  
 CC diversity of probes for measuring gene expression, with far less bias  
 CC than expressed sequence tag microarrays. The method is suitable for  
 CC rapid production of functional information from genomic sequence. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences.  
 XX  
 SQ Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;  
 Query Match 20.8%; Score 486; DB 22; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;  
 Matches 486; Conservative 0; Mismatches 0;  
 Oy 570 AAGTATGAATTTGACAGCAGCACTGACACCACTCCAGCGTGTGATGATGTAG 629  
 Db 486 AAGTATGAATTTGACAGCAGCACTGACACCACTCCAGCGTGTGATGATGTAG 427  
 Oy 630 AGTGATCAAG 689



04-FEB-2000; 2000US-0180312.  
 26-MAY-2000; 2000US-0207456.  
 30-JUN-2000; 2000US-0608408.  
 03-AUG-2000; 2000US-0632366.  
 21-SEP-2000; 2000US-0234687.  
 27-SEP-2000; 2000US-0236359.  
 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-488899/53.

Single exon nucleic acid probes for analyzing gene expression in human hearts -

Claim 1; SEQ ID No 809; 530bp; English.

The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.  
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPI at ftp.wpi.int/pub/published\_pct\_sequences.

Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;

Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

570 AAGTATGAATTTCGACGAGGAGTGAACACCACTCCGAGTGTGTCTATGTATGTAG 629

486 AAGTATGAATTTCGACGAGGAGTGAACACCACTCCGAGTGTGTCTATGTATGTAG 427

630 AGTATCAAG 689

426 AGTATCAAG 367

690 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCACCTGAGCTGAAGCGGACGGA 749

366 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCACCTGAGCTGAAGCGGACGGA 307

750 CGAGAGGAGCATTCGAAATCATCTACTACGCGAGAGATCTTTTACTGTGAGGTGGCTGC 809

306 CGAGAGGAGCATTCGAAATCATCTACTACGCGAGAGATCTTTTACTGTGAGGTGGCTGC 247

810 ACGACTTTCGAGAGTGGACCGAGATCGGGTGGCAGAAATCCAGTTCAATGTTGCT 869

246 ACGACTTTCGAGAGTGGACCGAGATCGGGTGGCAGAAATCCAGTTCAATGTTGCT 187

870 CAGAGAGATTCAGAGCGCGTGTCCCTGTTCTAATGCTGACACAGTACTGTTGATG 929

186 CAGAGAGATTCAGAGCGCGTGTCCCTGTTCTAATGCTGACACAGTACTGTTGATG 127

930 GCACCCGGAGATGACTTGGCCATCACTGAGTTTGTGTGATGCGACAGAGCATTTGG 989

126 GCACCCGGAGATGACTTGGCCATCACTGAGTTTGTGTGATGCGACAGAGCATTTGG 67

990 GACTGCTTTGAGAAAAAGATATATATAGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 1049

66 GACTGCTTTGAGAAAAAGATATATATAGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 7

1050 TGTCTG 1055

6 TGTCTG 1

RESULT 10

AAK00816/C

ID AAK00816 standard; DNA; 486 BP.

AAK00816;

05-NOV-2001 (first entry)

Human brain expressed single exon probe SEQ ID NO: 807.

Human; brain expressed exon; gene expression analysis; probe;

microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

epilepsy; cancer; 88.

Homo sapiens.

W0200157275-A2.

09-AUG-2001.

30-JAN-2001; 2001MO-US00667.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human brains -

Example 4; SEQ ID NO: 807; 650bp + Sequence listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.

Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;

Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

570 AAGTATGAATTTCGACGAGGAGTGAACACCACTCCGAGTGTGTCTATGTATGTAG 629

486 AAGTATGAATTTCGACGAGGAGTGAACACCACTCCGAGTGTGTCTATGTATGTAG 427

630 AGTATCAAG 689

426 AGTATCAAG 367

690 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCACCTGAGCTGAAGCGGACGGA 749

366 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCACCTGAGCTGAAGCGGACGGA 307

750 CGAGAGGAGCATTCGAAATCATCTACTACGCGAGAGATCTTTTACTGTGAGGTGGCTGC 809

306 CGAGAGGAGCATTCGAAATCATCTACTACGCGAGAGATCTTTTACTGTGAGGTGGCTGC 247



PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID No 836; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging  
 CC of diseases of the cervix, notably cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 AAGTATGAAATTCAGCGCAGCTGACACCACTCCAGCGTGTGCTCATGTATGTAG 629  
 DB 486 AAGTATGAAATTCAGCGCAGCTGACACCACTCCAGCGTGTGCTCATGTATGTAG 427  
 QY 630 ACTGATCAAG 689  
 DB 426 AGTATCAAG 367  
 QY 690 ATTATCCAGACCGAGCGCGAGTACAGCGCATCCAGCTCAGCTGAACTGGCAGCGGA 749  
 DB 366 ATTATCCAGACCGAGCGCGAGTACAGCGCATCCAGCTCAGCTGAACTGGCAGCGGA 307  
 QY 750 CGAGAGCGCATTCGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809  
 DB 306 CGAGAGCGCATTCGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247  
 QY 810 ACGACTCTTCGAGAGTGGAGAGCGGAGATGGGGTGGCAGAAATCCAGTTCATGTTGCT 869  
 DB 246 ACGACTCTTCGAGAGTGGAGAGCGGAGATGGGGTGGCAGAAATCCAGTTCATGTTGCT 187  
 QY 870 CAGAGAGAAATCAAGGCGGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTTCAATG 929  
 DB 186 CAGAGAGAAATCAAGGCGGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTTCAATG 127  
 QY 930 GCACCCGGGAATGACTTGGGCAATCACTGAGTTTGTGATGCGCAAGAGACATTTGG 989  
 DB 126 GCACCCGGGAATGACTTGGGCAATCACTGAGTTTGTGATGCGCAAGAGACATTTGG 67  
 QY 990 GACTGTCTTGAGAAACAGATATGATAGTGTGTTTACTGTTGTTCTTCTGAGATTG 1049  
 DB 66 GACTGTCTTGAGAAACAGATATGATAGTGTGTTTACTGTTGTTCTTCTGAGATTG 7  
 QY 1050 TGTCTG 1055  
 DB 6 TGTCTG 1

RESULT 13  
 AA132163/c  
 ID AA132163 standard; DNA; 486 BP.

XX AA132163:  
 AC 17-OCT-2001 (first entry)  
 XX DT  
 XX Probe #849 used to measure gene expression in human placenta sample.  
 DB  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KM genetic disorder; aa.  
 XX  
 OS Homo sapiens.

XX  
 PN MO200157272-A2.  
 XX

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0633366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 25; SEQ ID No 849; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs).

XX The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.

XX Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

XX Query Match 20.8%; Score 486; DB 22; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-129;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 AAGTATGAAATTCAGCGCAGCTGACACCACTCCAGCGTGTGCTCATGTATGTAG 629  
 DB 486 AAGTATGAAATTCAGCGCAGCTGACACCACTCCAGCGTGTGCTCATGTATGTAG 427  
 QY 630 ACTGATCAAG 689  
 DB 426 AGTATCAAG 367  
 QY 690 ATTATCCAGACCGAGCGCGAGTACAGCGCATCCAGCTCAGCTGAACTGGCAGCGGA 749  
 DB 366 ATTATCCAGACCGAGCGCGAGTACAGCGCATCCAGCTCAGCTGAACTGGCAGCGGA 307  
 QY 750 CGAGAGCGCATTCGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809  
 DB 306 CGAGAGCGCATTCGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247  
 QY 810 ACGACTCTTCGAGAGTGGAGAGCGGAGATGGGGTGGCAGAAATCCAGTTCATGTTGCT 869  
 DB 246 ACGACTCTTCGAGAGTGGAGAGCGGAGATGGGGTGGCAGAAATCCAGTTCATGTTGCT 187  
 QY 870 CAGAGAGAAATCAAGGCGGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTTCAATG 929  
 DB 186 CAGAGAGAAATCAAGGCGGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTTCAATG 127  
 QY 930 GCACCCGGGAATGACTTGGGCAATCACTGAGTTTGTGATGCGCAAGAGACATTTGG 989  
 DB 126 GCACCCGGGAATGACTTGGGCAATCACTGAGTTTGTGATGCGCAAGAGACATTTGG 67  
 QY 990 GACTGTCTTGAGAAACAGATATGATAGTGTGTTTACTGTTGTTCTTCTGAGATTG 1049  
 DB 66 GACTGTCTTGAGAAACAGATATGATAGTGTGTTTACTGTTGTTCTTCTGAGATTG 7  
 QY 1050 TGTCTG 1055  
 DB 6 TGTCTG 1

RESULT 14  
AAI00825/c  
ID AAI00825 standard; DNA; 486 BP.  
XX  
AC AAI00825;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Probe #816 used to measure gene expression in human breast sample.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder; ss;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 25; SEQ ID No 816; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes.  
CC The present sequence is one such probe. The probes are useful for  
CC measuring human gene expression in a human breast sample, where the probe  
CC hybridizes at high stringency to a nucleic acid expressed in the human  
CC breast. The probes are useful for predicting, diagnosing, grading,  
CC staging, monitoring and prognosing diseases of the human breast,  
CC particularly those diseases with polygenic aetiology. The diseases  
CC include: breast cancer, disorders of development, inflammatory diseases  
CC of the breast, fibrocystic changes, proliferative breast disease and  
CC non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;  
  
Query Match 20.8%; Score 486; DB 22; Length 486;  
Best Local Similarity 100.0%; Pred. No. 1.8e-129;  
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 570 AAGTATGAATTCACGCGAGTCTGACACCTCCAGCGTGTGTCTCATGTATGTGAG 629  
DB 486 AAGTATGAATTCACGCGAGTCTGACACCTCCAGCGTGTGTCTCATGTATGTGAG 427  
  
QY 630 AGTGATCAAGAGAAGGAGGAGAGGAGAAATGGAAGAATGAGGAGACCTAAGCCAAA 689  
DB 426 AGTGATCAAGAGAAGGAGGAGAGGAGAAATGGAAGAATGAGGAGACCTAAGCCAAA 367  
  
QY 690 ATTATCCAGACCGAGGCGGAGTACACCGCATCCCTCAGCTGAACTGGCACGCGA 749  
DB 366 ATTATCCAGACCGAGGCGGAGTACACCGCATCCCTCAGCTGAACTGGCACGCGA 307

QY 750 CGAGGACGCATTCCAAATATATCTACGGAGGAATCTTTTACTGTGAGGTGGCTGGTC 809  
DB 306 CGAGGACGCATTCCAAATATATCTACGGAGGAATCTTTTACTGTGAGGTGGCTGGTC 247  
QY 810 ACCACTTCTTCGGAGGTGGCAGCCGAGATCGGGTGGCAGAAATCCAGTTCAATGTGCT 869  
DB 246 ACCACTTCTTCGGAGGTGGCAGCCGAGATCGGGTGGCAGAAATCCAGTTCAATGTGCT 187  
QY 870 CAGAAGAGAAATCAAGCCCGTGTCCCTTCTTAATGCTGCACACCACTTACTTTCATG 929  
DB 186 CAGAAGAGAAATCAAGCCCGTGTCCCTTCTTAATGCTGCACACCACTTACTTTCATG 127  
QY 930 GCACCCGGGAATCACTTGGGCCAATCACTAGATTGTGTGTGTCGACAAAGACATTTGG 989  
DB 126 GCACCCGGGAATCACTTGGGCCAATCACTAGATTGTGTGTGTCGACAAAGACATTTGG 67  
QY 990 GACTGTCTTCAGAAAACAGATAATGATAGTCTTTTGTACTCTTCTTCTCGTAGGTTTC 1049  
DB 66 GACTGTCTTCAGAAAACAGATAATGATAGTCTTTTGTACTCTTCTTCTCGTAGGTTTC 7  
QY 1050 TGTCTG 1055  
DB 6 TGTCTG 1  
  
RESULT 15  
ABS00858/c  
ID ABS00858 standard; DNA; 486 BP.  
XX  
AC ABS00858;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human genome-derived single exon probe from lung SEQ ID No 849.  
XX  
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00665.  
XX  
PR 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2002-114183/15.  
XX  
PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
PS Claim 1; SEQ ID No 849; 634pp; English.  
XX







enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : [filang@lifetech.com](mailto:filang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 331 a 206 c 161 g 266 t 8 others  
ORIGIN

Query Match 40.3%; Score 939; DB 9; Length 972;  
Best Local Similarity 98.7%; Pred. No. 5,9e-237;

Matches 959; Conservative 7; Mismatches 4; Indels 2; Gaps 2;

1305 TTGTTTACATCTTTTGGAGCCAGAAAGCGTTTGGAAAAGTATCACTTTCCGAGA 1364  
1305 TTGTTTACATCTTTTGGAGCCAGAAAGCGTTTGGAAAAGTATCACTTTCCGAGA 1364  
972 TTTTATTACATCTTTTGGAGCCAGAAAGCGTTTGGAAAAGTATCACTTTCCGAGA 913  
1365 TTCTCGGATTTCTGACTCTTTTGCACAGCACTTCTGGGAACTCTTCTGGAAATGCA 1424  
1365 TTCTCGGATTTCTGACTCTTTTGCACAGCACTTCTGGGAACTCTTCTGGAAATGCA 1424  
912 TTCTCGGATTTCTGACTCTTTTGCACAGCACTTCTGGGAACTCTTCTGGAAATGCA 853  
912 TTCTCGGATTTCTGACTCTTTTGCACAGCACTTCTGGGAACTCTTCTGGAAATGCA 853  
1425 TCACCTAGAGATCCCAACCGTGCAGCTGTAACCTTGTGCTTTGCAAAAAGATGATCT 1484  
1425 TCACCTAGAGATCCCAACCGTGCAGCTGTAACCTTGTGCTTTGCAAAAAGATGATCT 1484  
852 TCACCTAGAGATCCCAACCGTGCAGCTGTAACCTTGTGCTTTGCAAAAAGATGATCT 793  
1485 GAAATTCCTCTGAGAAATTTAGCTTATACATTCAGAG-ATATGCACTTCACTGCGAAC 1543  
1485 GAAATTCCTCTGAGAAATTTAGCTTATACATTCAGAG-ATATGCACTTCACTGCGAAC 1543  
792 GAAATTCCTCTGAGAAATTTAGCTTATACATTCAGAG-ATATGCACTTCACTGCGAAC 733  
1544 TTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 1603  
1544 TTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 1603  
732 TTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 673  
1604 TCTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 1663  
1604 TCTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 1663  
672 TCTTTATGCGGTGAGAAATTTTATAGGTGTTGGGATCGGACCTCAGTTTCTGTGTT 613  
1664 GTTGAGATAGTGTCTTTTACCCCGAAGACATCAAGATACACTGTAATTAAGCT 1723  
1664 GTTGAGATAGTGTCTTTTACCCCGAAGACATCAAGATACACTGTAATTAAGCT 1723  
612 GTTGAGATAGTGTCTTTTACCCCGAAGACATCAAGATACACTGTAATTAAGCT 553  
1724 GATGCAATATATATACACTGTTTACACTGTTGGAAGTATGCGAGTGGAGACTAA 1783  
1724 GATGCAATATATATACACTGTTTACACTGTTGGAAGTATGCGAGTGGAGACTAA 1783  
552 GATGCAATATATATACACTGTTTACACTGTTGGAAGTATGCGAGTGGAGACTAA 493  
1784 GATGATTTTAACTTACCTGTGAAATCATATGTTTGAAGAAAGCTGTTCCATGTTAAACG 1843  
1784 GATGATTTTAACTTACCTGTGAAATCATATGTTTGAAGAAAGCTGTTCCATGTTAAACG 1843  
492 GATGATTTTAACTTACCTGTGAAATCATATGTTTGAAGAAAGCTGTTCCATGTTAAACG 433  
1844 GACTTGAATTTAAAGCAATGTAAGTAGATGTTGGCAGATGAGAGGAGTCA 1903  
1844 GACTTGAATTTAAAGCAATGTAAGTAGATGTTGGCAGATGAGAGGAGTCA 1903  
432 GACTTGAATTTAAAGCAATGTAAGTAGATGTTGGCAGATGAGAGGAGTCA 373  
1904 GTCCTTTCCCATTCATCTCTGATGAAATTTGTAATACAGTTTAACTTTGTAATTTT 1963  
1904 GTCCTTTCCCATTCATCTCTGATGAAATTTGTAATACAGTTTAACTTTGTAATTTT 1963  
372 GTCCTTTCCCATTCATCTCTGATGAAATTTGTAATACAGTTTAACTTTGTAATTTT 313  
1964 TTTAGTTGTAATGTAATGTCGTGTAATAGTAATTAATTTTGGCCTTAAACATACCGT 2023  
1964 TTTAGTTGTAATGTAATGTCGTGTAATAGTAATTAATTTTGGCCTTAAACATACCGT 2023  
312 TTTAGTTGTAATGTAATGTCGTGTAATAGTAATTAATTTTGGCCTTAAACATACCGT 253  
2024 AACATGTTTGCATTTTGAATATCTTAATGCAAGTAAACATGATGCTTTGGAAATTT 2083  
2024 AACATGTTTGCATTTTGAATATCTTAATGCAAGTAAACATGATGCTTTGGAAATTT 2083  
252 AACATGTTTGCATTTTGAATATCTTAATGCAAGTAAACATGATGCTTTGGAAATTT 193  
2084 GGAAGATGTTTATTTTGAAGCAAAATATGTTGCAATTAATG-CTTTGATTTTC 2142  
2084 GGAAGATGTTTATTTTGAAGCAAAATATGTTGCAATTAATG-CTTTGATTTTC 2142  
192 GGAAGATGTTTATTTTGAAGCAAAATATGTTGCAATTAATG-CTTTGATTTTC 133  
2143 ATATCAAGAAATTTGATGAAGCTTCAAAACCTGTTTACGTTAGTAAAGAGGAGC 2202

132 ATATCAAGAAATTTGATTAACGTTCTCAAAACCTGTTTACGTTAGTAAAGAGGAGC 73  
2203 CGGTTGGGAGAGACCATTCATCGCTGTCGAAGTGTCTTTGTAAGTCTTTTAACT 2262  
72 CGGTTGGGAGAGACCATTCATCGCTGTCGAAGTGTCTTTGTAAGTCTTTTAACT 13  
2263 GGAGAGCTTAAC 2274  
12 GGAGAGGTTAAC 1

RESULT 2  
AL546617 939 bp mRNA linear EST 16-FEB-2001

LOCUS AL546617 LTI.NFL006.PL2 Homo sapiens cDNA clone CS001029YJ21 5

DEFINITION prime, mRNA sequence.

ACCESSION AL546617 GI:12879906

VERSION AL546617.1 GI:12879906

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 939)

AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES

source

1. 939

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS01029YJ21"

/clone\_1bp="LTI.NFL006.PL2"

/issue="type="Placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : [filang@lifetech.com](mailto:filang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 263 a 219 c 238 g 218 t 1 others

ORIGIN

Query Match 40.3%; Score 938.6; DB 9; Length 939;  
Best Local Similarity 99.9%; Pred. No. 7.4e-237;  
Matches 938; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

92 ACTAGGGGCTTACATCGCGGCTGTGAGCGCTTCACTGTAAAGAACAGATGATTT 151  
92 ACTAGGGGCTTACATCGCGGCTGTGAGCGCTTCACTGTAAAGAACAGATGATTT 151  
1 ACTAGGGGCTTACATCGCGGCTGTGAGCGCTTCACTGTAAAGAACAGATGATTT 60  
152 TAGAACTTTAACTACAGTTTAACTCCCTGATTCCTGTTGCTGCAAAACAGTGAATCTT 211  
152 TAGAACTTTAACTACAGTTTAACTCCCTGATTCCTGTTGCTGCAAAACAGTGAATCTT 211  
61 TAGAACTTTAACTACAGTTTAACTCCCTGATTCCTGTTGCTGCAAAACAGTGAATCTT 120  
212 CAGCGAAAGTGAACCCAGGCGCAATTTGAGTCCCTCTTAGAGAGTATGACAGACAT 271  
212 CAGCGAAAGTGAACCCAGGCGCAATTTGAGTCCCTCTTAGAGAGTATGACAGACAT 271  
121 CAGCGAAAGTGAACCCAGGCGCAATTTGAGTCCCTCTTAGAGAGTATGACAGACAT 180  
272 CACCTTCAGTATTTTAAAGCTTCAAGAGTCAAGATTAACCTTCAAGAACCCCTTCTC 331  
272 CACCTTCAGTATTTTAAAGCTTCAAGAGTCAAGATTAACCTTCAAGAACCCCTTCTC 331  
181 CACCTTCAGTATTTTAAAGCTTCAAGAGTCAAGATTAACCTTCAAGAACCCCTTCTC 240  
332 CCGAGAGATGCCAGGCTCAGTGCATTAAGACTGAGTTTCTGGGAAAGAAATGAAGTT 391

```
Db 241 CGCAGCAGATGCCAGGCTCCAGCTGCATAGAGCTGTTCTGGGAAAGGAAATGAAGTT 300
QY 392 ATATTTTGTCTCAGACCTTACATAGGAAGCTCAGACCTGGCTCCGCCAAATCCAGACAA 451
Db 301 ATATTTTGTCTCAGACCTTACATAGGAAGCTCAGACCTGGCTCCGCCAAATCCAGACAA 360
QY 452 GCAGTTTCTGATCTCCCTCCGCTCTCCGCTGAGTGGATGGAAACAAAGTGGAGATGC 511
Db 361 GCAGTTTCTGATCTCCCTCCGCTCTCCGCTGAGTGGATGGAAACAAAGTGGAGATGC 420
QY 512 GACCCGAGTCATAAATCTGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAA 571
Db 421 GACCCGAGTCATAAATCTGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAA 480
QY 572 GTATGAATGTCAGCAGGAGTACACCACTCCAGCGTGGTGTCTCATGTATGTGAGAG 631
Db 481 GTATGAATGTCAGCAGGAGTACACCACTCCAGCGTGGTGTCTCATGTATGTGAGAG 540
QY 632 TGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 691
Db 541 TGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 692 TATCAGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
Db 601 TATCAGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 752 AGGAGGAGTCCAAATCTATCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
Db 661 AGGAGGAGTCCAAATCTATCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 812 GACTTCTTCGAGGTGGCAGCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 871
Db 721 GACTTCTTCGAGGTGGCAGCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 780
QY 872 GAAGAGATCAAGCGGTGCTCCCTGTTCTAATGCTGCACACCAAGTACTGTTTCTATGGC 931
Db 781 GAAGAGATCAAGCGGTGCTCCCTGTTCTAATGCTGCACACCAAGTACTGTTTCTATGGC 840
QY 932 ACCCGGAATGACTTGGGCGCAATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 991
Db 841 ACCCGGAATGACTTGGGCGCAATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 900
QY 992 CTGTCTTGAGAAACAGATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1030
Db 901 CTGTCTTGAGAAACAGATATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 939

RESULT 3
AL567504/c
LOCUS AL567504 LTI_FL013 FBrn1 Homo sapiens cDNA clone CS0DF037YL22 3
DEFINITION prime, mRNA sequence.
ACCESSION AL567504
VERSION AL567504.1 GI:12920929
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..959
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF037YL22"
/clone_lib="LTI_FL013_FBrn1"
```

/dev\_stages="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DH10B"  
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 329 a 193 c 162 g 263 t 12 others

Query Match 39.4%; Score 919; DB 9; Length 959;  
Best Local Similarity 97.8%; Pred. No. 1.1e-231;  
Matches 938; Conservative 12; Mismatches 7; Indels 2; Gaps 2;

```
QY 1356 TTTCCAGATTCTCGAATCTCGACTCTTTGCAACAGCACTTGTCTGGGAACCTTCTCT 1415
Db 959 TTTCCAGATTCTCGAATCTCGACTCTTTGCAACAGCACTTGTCTGGGAACCTTCTCT 900
QY 1416 GGAATGCATTCACTCAGCAT-CCCCAACCGTGCAACGTGTAACCTTGTCTTTCGAAAG 1474
Db 899 GGAATGCATTCACTCAGCATCCCCAACCGTGCAACGTGTAACCTTGTCTTTCGAAAG 840
QY 1475 AAGTTGATCTGAAATTCCTCTGTAGAATTTAGCTTATACAATTCAGAGATAGCAGTTTC 1534
Db 839 AAGTTGATCTGAAATTCCTCTGTAGAATTTAGCTTATACAATTCAGAGATAGSAG-TTC 781
QY 1535 ACTGCCAACTTTTGTAGTGGTGAGAAATTTAGTTAGTGTGTGGGATCGGACCTCAGTT 1594
Db 780 ACTGCCAACTTTTGTAGTGGTGAGAAATTTTGTAGTTAGTGTGTGGGATCGGACCTCAGT 721
QY 1595 TCTGTGTCTTTTATGTGTGTGTGTCTATACATGAATCATAGCAACCAAAACCTTTTGG 1654
Db 720 TCTGTGTGTCTTTTATGTGTGTGTGTCTATACATGAATCATAGCAACCAAAACCTTTTGG 661
QY 1655 AAACGTGTGTGTAGATGTGTGTCTTTTACCCAGAAAGACATCAAGATACACATTGTA 1714
Db 660 AAACGTGTGTGTAGATGTGTGTCTTTTACCCAGAAAGACATCAAGATACACATTGTA 601
QY 1715 AATAAAGCTGATAGCATATATTCATACCTCTGTGTACACTTGGGTGAAAGTATGGCAGTG 1774
Db 600 AATAAAGCTGATAGCATATATTCATACCTCTGTGTACACTTGGGTGAAAGTATGGCAGTG 541
QY 1775 GGAGACTAAGATGATTAACCTACCTGTGAATCATATGTTGTAGGAAAGCTGTCTCCCAT 1834
Db 540 GGAGACTAAGATGATTAACCTACCTGTGAATCATATGTTGTAGGAAAGCTGTCTCCCAT 481
QY 1835 GTCTAACAGGACTTGAATTCAAAGCATGTCAAGTGAATAGTATCTGTGGCGATATGAG 1894
Db 480 GTCTAACAGGACTTGAATTCAAAGCATGTCAAGTGAATAGTATCTGTGGCGATATGAG 421
QY 1895 AGGATGCACTGCTCTTCCCATTCATCTCTGTGAATGTTGTATAGTATTAATTTTGGCCTTA 1954
Db 420 AGGATGCACTGCTCTTCCCATTCATCTCTGTGAATGTTGTATAGTATTAATTTTGGCCTTA 361
QY 1955 GTAATTTTTTCTAGTGTGAATGTGTATGTCTGGTAAATAGTATTAATTTTGGCCTTA 2014
Db 360 GTAATTTTTTCTAGTGTGAATGTGTATGTCTGGTAAATAGTATTAATTTTGGCCTTA 301
QY 2015 CAATACCGTAACAATGTTGTCAATTTTGAATTAATTAATGCAAGTAAAGTATGCTT 2074
Db 300 CAATACCGTAACAATGTTGTCAATTTTGAATTAATTAATGCAAGTAAAGTATGCTT 241
QY 2075 TGGAAATTTGGAAGTGTGTTTATTTCTTTGAGAAGCAATATGTTTGTGATTAATGCTTT 2134
Db 240 TGGAAATTTGGAAGTGTGTTTATTTCTTTTGAAGCAATATGTTTGTGATTAATGCTTT 181
QY 2135 GATTGTTTCATATCAAGAAATTCGATGAAACCTTCTCAAAACCTTGTTCGGTACTTGGTAA 2194
```

```

|||||
Db 180 GATTTCATATCAAGAAATTGATGAAGTTCACAAACCCGTTTACGTAATCTTGTA 121
Oy 2195 GAGGAGAGCCGCTTTGGAGAGACCAATGATGCTGCTCCAGATGTTTCTGTAAGTCT 2254
Db 120 GAGGAGAGCCGCTTTGGAGAGACCAATGATGCTGCTCCAGATGTTTCTGTAAGTCT 61
Oy 2255 TTTAACTGAGAGAGCTTCAAAATATCTTTTAACTGATCTTATATTAATG 2313
Db 60 TTTAACTGAGAGAGCTTCAAAATATCTTTTAACTGATCTTATATTAATG 2

RESULT 4
AL543576 931 bp mRNA linear EST 16-FEB-2001
LOCUS AL543576 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D106YK09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543576 GI:12876055
VERSION AL543576
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 931
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D106YK09"
/clone_1lb="LTI_NFL006.PL2"
/clone_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t

Query Match 39.3%; Score 915.4; DB 9; Length 931;
Best local similarity 99.8%; Pred. No. 1e-230; Indels 1; Gaps 1;
Matches 927; Conservative 0; Mismatches 1;

Oy 95 AGGGCTTACTGCTGGTGTGTAGCCCTTCACTGTAGAAAGCAAGATGCAATTTAG 154
Db 1 AGGGCTTACTGCTGGTGTGTAGCCCTTCACTGTAGAAAGCAAGATGCAATTTAG 60
Oy 155 AAACCTTAACTAAGCTTTAGCTCCCTGATGCTGTGTGGCAACAGTATATCTTCA 214
Db 61 AAACCTTAACTAAGCTTTAGCTCCCTGATGCTGTGTGGCAACAGTATATCTTCA 120
Oy 215 CGAAAGTGAACCAAGGCCCAATTTGAGCCCTCTTTAGAGAGTATGACAGACATAC 274
Db 121 CGAAAGTGAACCAAGGCCCAATTTGAGCCCTCTTTAGAGAGTATGACAGACATAC 180
Oy 275 CTTTCACTATTTTAAAGAGCTTCAAGAGTCAAGATTAATCTTCAAGCAACCCCTTCCGC 334
Db 181 CTTTCACTATTTTAAAGAGCTTCAAGAGTCAAGATTAATCTTCAAGCAACCCCTTCCGC 240
Oy 335 AGCAGATGCCAGGCTCCAGCTGCAATGACTGATTTCTGGAGAAAGAAATGAATTATA 394

```

```

|||||
Db 241 AGCAGATGCCAGGCTCCAGCTGCAATGACTGATGATTTCTGGAGAAAGAAATGAATTATA 300
Oy 395 TTTTGTCTAGACCTTACATATGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAGCA 454
Db 301 TTTTGTCTAGACCTTACATATGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAGCA 360
Oy 455 GTTTCTGATCTCCCTCCGCTCCGCTCTCTCCGCTGAGTGGATGAAACAAAGTGAATGCGAC 514
Db 361 GTTTCTGATCTCCCTCCGCTCCGCTCTCTCCGCTGAGTGGATGAAACAAAGTGAATGCGAC 420
Oy 515 CCCAGTATTAATCTATGATCTCTTATATGCAATCTCCAGAGCTGGGGCCAGGGGAAAGTA 574
Db 421 CCCAGTATTAATCTATGATCTCTTATATGCAATCTCCAGAGCTGGGGCCAGGGGAAAGTA 480
Oy 575 TGAATTGACAGAGCCGACTGACACCACTCCACAGCTGTGTGTCATGTATGTGAGAGTGA 634
Db 481 TGAATTGACAGAGCCGACTGACACCACTCCACAGCTGTGTGTCATGTATGTGAGAGTGA 540
Oy 635 TCAGAGAGAGAGAGAGAGAGAGAAATGGAAGATGAGAGACCTTAAGCCAAAATTTAT 694
Db 541 TCAGAGAGAGAGAGAGAGAGAGAAATGGAAGATGAGAGACCTTAAGCCAAAATTTAT 600
Oy 695 CCAGACCAAGAGAGCCGAGATGACAGCCGATCCACTGAGTGAATGACAGCAGAGAG 754
Db 601 CCAGACCAAGAGAGCCGAGATGACAGCCGATCCACTGAGTGAATGACAGCAGAGAG 660
Oy 755 AGCATTCCAATCATCTACACGGAGAGAACTTTTATCTGTGAGAGTGTGCTGTACAGAC 814
Db 661 AGCATTCCAATCATCTACACGGAGAGAACTTTTATCTGTGAGAGTGTGCTGTACAGAC 720
Oy 815 TTTCTGAGAGTGGACAGCCGATGCGGGGTGGAGAAATCCAGTTATATGTTGCTCGA 874
Db 721 TTTCTGAGAGTGGACAGCCGATGCGGGGTGGAGAAATCCAGTTATATGTTGCTCGA 780
Oy 875 GAGATCAAGGCCGCTGCTCCCTGTTCTAATGTCTGACACACAGTACTGTTCATGAGCACC 934
Db 781 GAGATCAAGGCCGCTGCTCCCTGTTCTAATGTCTGACACACAGTACTGTTCATGAGCACC 840
Oy 935 CGGGAATGACTTGGCCCAATCATGAGTGTGTGTGATGCGACAAGACATTTGGAGCTG 994
Db 841 CGGGAATGACTTGGCCCAATCATGAGTGTGTGTGATGCGACAAGACATTTGGAGCTG 900
Oy 995 TCTTGAGAAACAGTAAATGATAGCTT 1023
Db 901 TCTTGAG-AAACAGATAAATGATAGTGT 928

RESULT 5
AL554686 939 bp mRNA linear EST 16-FEB-2001
LOCUS AL554686 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1085YF09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL554686
VERSION AL554686 GI:12895707
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 939
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1085YF09"

```

```

/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      262 a  219 c  238 g  218 t      2 others
ORIGIN

Query Match      39.0%; Score 908.4; DB 9; Length 939;
Best Local Similarity 99.3%; Pred. No. 7.1e-229;
Matches 933; Conservative 1; Mismatches 3; Indels 3; Gaps 2;

QY  92  ACTAGGGGCTTCACTCGCTGGGTCTGTAGCGCTTTACCTGTAAAGAACCAAGATGCAATTT 151
Db  1  ACTAGGGGCTTCACTCGCTGGGTCTGTAGCGCTTTACCTGTAAAGAACCAAGATGCAATTT 60

QY  152  TAGAAACTTTAACTACAGTTTTAGCTCCCTGATTCCTGCTGCTGCTGCAACACAGTATATCTT 211
Db  61  TAGAAACTTTAACTACAGTTTTAGCTCCCTGATTCCTGCTGCTGCTGCAACACAGTATATCTT 120

QY  212  CAGCCAAAGTGAACCA--GGCCAAATTTGAGTCCCTCTTTAGGACGTATGCAAGGAC 269
Db  121  CAGCCAAAGTGAACCAATGTGCCAAATTTGAGTCCCTCTTTAGGACGTATGCAAGGAC 180

QY  270  ATCACTTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGATTAACATTCAGAACCCCTTC 329
Db  181  ATCACTTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGATTAACATTCAGAACCCCTTC 240

QY  330  TCCGACGAGATGCCAGCTCCAGCTCATAGAAGTCTGAGTTTCTGGAAAGGAATGAAG 399
Db  241  TCCGACGAGATGCCAGCTCCAGCTCATAGAAGTCTGAGTTTCTGGAAAGGAATGAAG 300

QY  390  TTATATTTTGTCTCAGACTTACATAGGAGCTCACACCTGGCTCGCCAAATCCAGAC 449
Db  301  TTATATTTTGTCTCAGACTTACATAGGAGCTCACACCTGGCTCGCCAAATCCAGAC 360

QY  450  AAGCAGTTTCTGATTCCTCCCTCCGCTCTCCGCTAGTGGATGGAACAGAGTGAAGAT 509
Db  361  AAGCAGTTTCTGATTCCTCCCTCCGCTCTCCGCTAGTGGATGGAACAGAGTGAAGAT 420

QY  510  GCGACCCAGCTCAATACTATGATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGAA 569
Db  421  GCGACCCAGCTCAATACTATGATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGAA 480

QY  570  AAGTATGAATTCGACGAGGAGCTGACACCACTCCGAGGCTGGTGTCTCATGTATGTAG 629
Db  481  AAGTATGAATTCGACGAGGAGCTGACACCACTCCGAGGCTGGTGTCTCATGTATGTAG 540

QY  630  AGTGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 689
Db  541  AGTGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600

QY  690  ATTATCCAGACAGGAGGCGGAGTACAGCCGATCCACTCAGCTGAACTGGCAGCGGA 749
Db  601  ATTATCCAGACAGGAGGCGGAGTACAGCCGATCCACTCAGCTGAACTGGCAGCGGA 660

QY  750  CGAGGACGCAATCCAAATCATACTCAGCGGAGGAGTCTTTTCTGAGGAGTGGCTGTC 809
Db  661  CGAGGACGCAATCCAAATCATACTCAGCGGAGGAGTCTTTTCTGAGGAGTGGCTGTC 720

QY  810  ACCACTTCTTCGAGGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAATTCATGTCCT 869
Db  721  ACCACTTCTTCGAGGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAATTCATGTCCT 780

QY  870  CAGAAGAGAAATCAAGGCGGTGTCCCTTGTCTTAATGCTGCACACCACTGCTGTCATG 929

```

```

Db  781  CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTTAATGCTGCACACCACTTACTGTTCATG 840
QY  930  GCACCGGGGAATGACTTGGCCCAATCACTGAGTTGTTGTTGATCGCACAGGACATTTGG 989
Db  841  GCACCGGGGAATGACTTGGCCCAATCACTGAGTTGTTGTTGATCGCACAGGACATTTGG 900
QY  990  GACTGTCTTTCAGAAACAGATAATGATGATGTTTGTACT 1029
Db  901  GACTGTCTTTCAGAAACAGATAATGATGATGTTTGTACT 939

RESULT 6
AL544733/c
LOCUS      AL544733      938 bp      mRNA      linear      EST 16-FEB-2001
DEFINITION AL544733 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1022VJ23 3
prime, mRNA sequence.
ACCESSION  AL544733
VERSION    AL544733.1 GI:12877213
KEYWORDS   EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 938)
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             source
                     1..938
                        Location/Qualifiers
                        1..938
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="CS0D1022VJ23"
                        /clone_lib="LTI_NFL006_PL2"
                        /tissue_type="placenta"
                        /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
                        was primed with a NotI-oligo(dT) primer. Five prime end
                        enriched, double-stranded cDNA was digested with Not I and
                        cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                        vector. Library was normalized. Library was constructed by
                        Life Technologies. Contact : Feng Liang Life Technologies,
                        a division of Invitrogen 9800 Medical Center Drive
                        Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                        Email : fliang@lifetech.com URL :
                        http://fulllength.invitrogen.com"
BASE COUNT      325 a  194 c  153 g  263 t      3 others
ORIGIN

Query Match      38.6%; Score 900.6; DB 9; Length 938;
Best Local Similarity 98.9%; Pred. No. 8.2e-227;
Matches 914; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY  1371  GATTCCTGACTCTTTGCAACAGACACTTGTCTGGGAACTCTCTCTGGAATGCATTCACTC 1430
Db  930  GAATCTCRACTTTTACACAGAACTTGTCTGGAACTCTCTCTGGAATGCATTCACTC 871

QY  1431  AGCATCCCCAACCGTGCACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAT 1490
Db  870  AGCATCCCCAACCGTGCACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAT 811

QY  1491  CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATGAGCTTCACCTGCCAATTTAGT 1550
Db  810  CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATGAGCTTCACCTGCCAATTTAGT 752

QY  1551  GGTGTAGAAATTTAGTTTAGTGTGGATCGGACCTCAGTTCCTGTTGTTCTTTTA 1610
Db  751  GGTGTAGAAATTTAGTTTAGTGTGGATCGGACCTCAGTTCCTGTTGTTCTTTTA 692

QY  1611  TGTGTGTTTCTTATACATGAATCATAGCCAAATCTTTTGGAAACTGTTGTTGAGA 1670

```



```

QY 1795 CTACCTGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGTTCTAACAGGACTTGAATTC 1854
Db 841 CTACCTGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGTTCTAACAGGACTTGAATTC 900
QY 1855 AAAGCATGTCAAGT 1868
Db 901 AAAGCATGTTAAGT 914

RESULT 8
BM450020
LOCUS BM450020
DEFINITION AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
5', mRNA sequence.
ACCESSION BM450020
VERSION BM450020.1 GI:18499060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1041)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12204 row: d column: 10
High quality sequence stop: 643.
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528025"
/clone_lib="NIH MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 282 a 248 c 274 g 236 t 1 others
ORIGIN

Query Match 37.3%; Score 869.6; DB 13; Length 1041;
Best Local Similarity 98.7%; Pred. No. 1.3e-218;
Matches 898; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 61 AGGAACCTACAGCCTCTTGGAAAGGAATCTCACTAGGGCTTGACCTGGCTGTGTAG 120
Db 1 AGGAACCTACAGCCTCTTGGAAAGGAATCTCACTAGGGCTTGACCTGGCTGTGTAG 60
QY 121 CGCTTTCACTGTGAAGCAAGATGCAATTTAGAACTTTAACTACAGTTTAGCTCCC 180
Db 61 CGCTTTCACTGTGAAGCAAGATGCAATTTAGAACTTTAACTACAGTTTAGCTCCC 120
QY 181 TGAATGCTGTGTGGCAACAGTGATCTTCAGCGAAAGTGAACACCGGCGCAATTTG 240
Db 121 TGAATGCTGTGTGGCAACAGTGATCTTCAGCGAAAGTGAACACCGGCGCAATTTG 180
QY 241 AGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTAAGACTTCAAC 300
Db 181 AGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTAAGACTTCAAC 240
QY 301 GAGTCAGATAAATCTCAGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTGCATA 360

```

```

Db 241 GAGTCAGATAAATCTCAGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTGCATA 300
QY 361 AGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGTCTAGACCTTACACATAGGAA 420
Db 301 AGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGTCTAGACCTTACACATAGGAA 360
QY 421 GCTCACACCTGCTCCGCCAAATCCAGACAAGCAAGTTTCTGATCTCCCTCCCGCTCTC 480
Db 361 GCTCACACCTGCTCCGCCAAATCCAGACAAGCAAGTTTCTGATCTCCCTCCCGCTCTC 420
QY 481 CGCCAGTGGGATGGAACCAAGTGAAGATGCCACCCAGTCAATAACTATGATCTCTTAT 540
Db 421 CGCCAGTGGGATGGAACCAAGTGAAGATGCCACCCAGTCAATAACTATGATCTCTTAT 480
QY 541 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTCACCGCAGCTGACACCA 600
Db 481 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTCACCGCAGCTGACACCA 540
QY 601 CTCCAGCTGGTGGTCCATGATGTGAGAGTGATCAAGAGAAGGAGGAAGAGGAAA 660
Db 541 CTCCAGCTGGTGGTCCATGATGTGAGAGTGATCAAGAGAAGGAGGAAGAGGAAA 600
QY 661 TGGAAAGAAATGAGGAGACCTTAAGCCCAAAATTTATCCAGCAGGAGGCGGAGTACAGC 720
Db 601 TGGAAAGAAATGAGGAGACCTTAAGCCCAAAATTTATCCAGCAGGAGGCGGAGTACAGC 660
QY 721 CGATCCACTCAGCTGAATCGCAGCGACGAGGACGATTCMAATCATATCACTACGGGA 780
Db 661 CGATCCACTCAGCTGAATCGCAGCGACGAGGAGGATTCMAATCATATCACTACGGGA 720
QY 781 GGAATCTTTTACTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840
Db 721 GGAATCTTTTACTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
QY 841 GGGTGGCAAAATCCAGTTCATGTTGCTCAGAAAGAGATCAAGGCCGCTGTCCTTTGTT 900
Db 781 GGGTGGCAAAATCCAGTTCATGTTGCTCAGAAAGAGATCAAGGCCGCTGTCCTTTGTT 840
QY 901 CTAATGCTGCACA-CCAGTTACTGTTCAATGCAACCCGGG--AATGACTTGGGCCAATCAC 957
Db 841 CTAATGCTGCACACCCAGTTACTGTTCAATGCAACCCGGGAAATGACTTGGGCCAATCAC 900
QY 958 TCAGTTTGTG 967
Db 901 TGGAGTTTG 910

RESULT 9
AL544313
LOCUS AL544313
DEFINITION AL544313 LTI_NFL006_P12 Homo sapiens cDNA clone CSODI019YH04 5
prime, mRNA sequence.
ACCESSION AL544313
VERSION AL544313.1 GI:12876793
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 906)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequençage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI019YH04"

FEATURES
source

```



```

/clone_lib="LTI_NFL006_PL2"
/issue_type="placenta"
/notes="Vector: PCWVSPO8 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPO8 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 253 a 215 c 229 g 202 t 7 others
ORIGIN

```

```

Query Match 37.1%; Score 864; DB 9; Length 906;
Best Local Similarity 98.7%; Pred. No. 3.9e-217;
Matches 894; Conservative 2; Indels 3; Gaps 3;

OY 100 CTGACTGCGTGGTCTGTAGCCCTTCACTGTAGAAAGCAAGATGATTTAGAACT 159
DB 1 CTGACTGCGTGGTCTGTAGCCCTTCACTGTAGAAAGCAAGATGATTTAGAACT 60
OY 160 TTACTACAGTTTATAGTCCCTGATTGCTGTGTGGCAACAGTATATCTTCAAGGAAA 219
DB 61 TTACTACAGTTTATAGTCCCTGATTGCTGTGTGGCAACAGTATATCTTCAAGGAAA 120
OY 220 GTGAAACCAAGGCGCAATTTAGTCCCTCTTTAGACATGTAGCAAGGACATCACTTTC 279
DB 121 GTGAAACCAAGGCGCAATTTAGTCCCTCTTTAGACATGTAGCAAGGACATCACTTTC 180
OY 280 AGATTTTAAAGCTTCAACAGAGTCAAGTAATACTTCAAGCAACCCCTTCTCCGACAG 339
DB 181 AGATTTTAAAGCTTCAACAGAGTCAAGTAATACTTCAAGCAACCCCTTCTCCGACAG 240
OY 340 ATGCCAGGCTCCAGCTGCATTAAGTCTGTGTGGAAAGAAATGAATTTATTTTG 399
DB 241 ATGCCAGGCTCCAGCTGCATTAAGTCTGTGTGGAAAGAAATGAATTTATTTTG 300
OY 400 CTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAATCCAGCAACAGTTTC 459
DB 301 CTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAATCCAGCAACAGTTTC 360
OY 460 TGATCTCCCTCCGCTCTCCGCACTGAGTGAAGAAAGATGCAAGATGCAAGATGCAAG 519
DB 361 TGATCTCCCTCCGCTCTCCGCACTGAGTGAAGAAAGATGCAAGATGCAAGATGCAAG 420
OY 520 TCATTAATATGATCTTTATATGATCTTCAAGCTGGGCGCCAGGGAAGAAATGAAT 579
DB 421 TCATTAATATGATCTTTATATGATCTTCAAGCTGGGCGCCAGGGAAGAAATGAAT 480
OY 580 TGCAGCGCAGTACGACCACTCCAGCTGTGTGTGATGATGATGATGATGATGATGATG 639
DB 481 TGCAGCGCAGTACGACCACTCCAGCTGTGTGTGATGATGATGATGATGATGATGATG 539
OY 640 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 699
DB 540 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 599
OY 700 CCAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 759
DB 600 CCAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 659
OY 760 TTCCAATATATCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 819
DB 660 TTCCAATATATCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 719
OY 820 CGAGAGTGGCAGCCAGAGTGGGGTGGCAGAAATCCAGTTTATTTGCTCAGAAAGAAA 879
DB 720 CGAGAGTGGCAGCCAGAGTGGGGTGGCAGAAATCCAGTTTATTTGCTCAGAAAGAAA 779
OY 880 TCAGAGGCGTGTCTCTTTTCTAATGCTGACACAGATTAATGCTGACACCGGAGG 939

```

```

DB 780 TCAGAGGCGTGTCTCTTTTCTAATGCTGACACAGTACTGTTCAAGCA-CCGGGA 838
OY 940 ATGACTGGGCGCAATCACTAGTTTGTGTGATTCGACAGAGCAATTTGGACTGTCTTG 999
DB 839 ATGACTGGGCGCAATCACTAGTTTGTGTGATTCGACAGAGCA-TTGGGACTGTCTTG 897
OY 1000 AGAAAA 1005
DB 898 AGAAAA 903

```

```

RESULT 10
AL559594 1014 bp mRNA linear EST 16-FEB-2001
LOCUS AL559594 LTI_NFL008.TC2 Homo sapiens cDNA clone CS00D013YN24 5
DEFINITION
ACCESSION AL559594
VERSION AL559594.1 GI:12905231
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
source
1..1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS00D013YN24"
/clone_lib="LTI_NFL008.TC2"
/sex="male"
/issue_type="T cells from T cell leukemia"
/notes="Vector: PCWVSPO8 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPO8 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 271 a 237 c 270 g 234 t 2 others
ORIGIN

```

```

Query Match 36.7%; Score 855.2; DB 9; Length 1014;
Best Local Similarity 96.7%; Pred. No. 8.5e-215;
Matches 913; Conservative 2; Mismatches 25; Indels 4; Gaps 4;

OY 196 CAAACAGTATATCTTCAAGCAAGTGAAGCAAGGCGCCAAATTTGAGTCCCTTTAGCA 255
DB 51 CAATCTGCACTATTGGAAGAAAGCAATGTTGGCCAAATTTGAGTCCCTTTAGCA 110
OY 256 CGTATGACAGGACATCACTTTCAGTATTTTAAAGCTTCAAGAGTCAAGATGAATAACT 315
DB 111 CGTATGACAGGACATCACTTTCAGTATTTTAAAGCTTCAAGAGTCAAGATGAATAACT 170
OY 316 TCAGCAACCCCTTCTCCGAGAGAGTCCAGGCTCAGCTGCATTAAGACTGAGTTCTGG 375
DB 171 TCAGCAACCCCTTCTCCGAGAGAGTCCAGGCTCAGCTGCATTAAGACTGAGTTCTGG 230
OY 376 GAAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTGGCTC 435
DB 231 GAAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTGGCTC 290
OY 436 CGCCAAATCCAGCAAGCAAGTCTGATCTCCCTCCGCTCTCCGCGCAGTGGAGTGA 495

```



```

Db 291 CGCCAAATCCAGACAGAGTTCTGATCTCCCTCCGCTCTCCGCCAGTGGATGA 350
Qy 496 AACAAATGGAAGATGCGACCCCACTATAAATATGATCTCTTATATGCCATCTCAAGC 555
Db 351 AACAAATGGAAGATGCGACCCCACTATAAATATGATCTCTTATATGCCATCTCAAGC 410
Qy 556 TGGGGCCAGGGGAAAGATGATGAATTCAGCAGGACTGACACCACTCCAGCGTGGTG 615
Db 411 TGGGGCCAGGGGAAAGATGATGAATTCAGCAGGACTGACACCACTCCAGCGTGGTG 470
Qy 616 TCCATGATGATGAGAGTATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675
Db 471 TCCATGATGATGAGAGTATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 530
Qy 676 GACCTAAGCCAAATATTCAGACAGGAGGCGGAGTACAGCGGATCCACCTCAGCT 735
Db 531 GACCTAAGCCAAATATTCAGACAGGAGGCGGAGTACAGCGGATCCACCTCAGCT 590
Qy 736 GAACTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 795
Db 591 GAACTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 650
Qy 796 GAGGTGGCTGCTGACGACTTCTCGAGGTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAG 855
Db 651 GAGGTGGCTGCTGACGACTTCTCGAGGTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAG 710
Qy 856 CAGTTCATGCTCAGAGAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGCACACC 915
Db 711 CAGTTCATGCTCAGAGAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGCACACC 770
Qy 916 AGTTACTGTTTACGACCGGAGGAGTGGGCGCAATCACTGAGTGGTGGTGGTGGTGGT 975
Db 771 AGTTACTGTTTACGACCGGAGGAGTGGGCGCAATCACTGAGTGGTGGTGGTGGTGGT 830
Qy 976 ACAAGCAGATTTGGGAGTGTCTTGAGAAAAGATGATGATGATGATGATGATGATGAT 1035
Db 831 ACAAGCAGATTTGGGAGTGTCTTGAG-AAACAGATATGATGATGATGATGATGATGAT 889
Qy 1036 TTTCTGGTAGGTTCTGCTGTCGCAAGGCGAGTTGATCAGTGGTGGTGGTGGTGGTGGT 1095
Db 890 TTTCTGGTAGGTTCTGCTGTCGTCGCG-AGGCGGTTGATCAGTGGTGGTGGTGGTGGT 947
Qy 1096 CCTGTTTCTAAGTGGCTGCGAGGCGCACTCTCTACTGTTAGGA 1139
Db 948 CCTGTTTCTAAGTGGCTGCGAGGCG-CMCTCTCTACTGTTAGGA 990

RESULT 11
AL543715/c
LOCUS
DEFINITION
AL543715 LTI_NFL006_P12 Homo sapiens cdna clone EST 16-FEB-2001
prime, mRNA sequence.
ACCESSION
AL543715
VERSION
AL543715.1 GI:12876194
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 934)
Li, W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..934
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1006YK09"

```

```

/clone lib="LTI_NFL006_P12"
/tissue type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 308 a 194 c 156 g 260 t 16 others
ORIGIN

```

```

Query Match 36.1%; Score 841.8; DB 9; Length 934;
Best Local Similarity 96.0%; Pred. No. 2.9e-211;
Matches 872; Conservative 6; Mismatches 28; Indels 2; Gaps 2;
Qy 1355 CTTTCCAGATTCGGATTCGACTTTTGCACAGCACTTGTTCGGAACTCTTCC 1414
Db 906 CTTTCCAGATTCGCTGATTCGACTTTTGCACAGCACTTGTTCGGAACTCTTCC 848
Qy 1415 TGAATGCAATTCACCTCAGCATCCCAACCGTCAACGCTGTAACCTTGTGTTTGCAGAAAG 1474
Db 847 TGAATGCAATTCACCTCAGCATCCCAACCGTCAACGCTGTAACCTTGTGTTTGCAGAAAG 789
Qy 1475 AAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCACTTTC 1534
Db 788 AAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCACTTTC 729
Qy 1535 ACTGCCAATTTAGTGGGTGAGAAATTTTGTAGTGTTCGGATTCGGACCTCAGTT 1594
Db 728 ACGGCAATTTAGTGGGTGAGAAATTTTGTAGTGTTCGGATTCGGACCTCAGTT 669
Qy 1595 TCTGTTGTTCTTTTATGTTGTTTCTATACATGATCATGCAAAATCTTTTGG 1654
Db 668 TCTGNNGTTCCTTTTATGTTGTTTCTATACATGATCATGCAAAATCTTTTGG 609
Qy 1655 AAACCTGTTGTTGAGATGTTGTTTCTTTTACCCAGAGACATCAAGATACACTTGTGA 1714
Db 608 AAACCTGTTGTTGAGATGTTGTTTCTTTTACCCAGAGACATCAAGATACACTTGTGA 549
Qy 1715 AATAAGCTGATAGCATATATTACATCTGTTGACACTTGGGTGAAAGATATGGCAGTG 1774
Db 548 AATAAGCTGATAGCATATATTACATCTGTTGACACTTGGGTGAAAGATATGGCAGTG 489
Qy 1775 GGAGACTAAGATGATTAACTACCTGTAATCATATGTTGTAGGAAAGCTGTTCCTCAT 1834
Db 488 GGAGACTAAGATGATTAACTACCTGTAATCATATGTTGTAGGAAAGCTGTTCCTCAT 429
Qy 1835 GTCTAACAGGACTTGAATTCAGAGCATGTCAGTGGATAGTAGATCTGTGGCGATATGAG 1894
Db 428 GTCTAACAGGACTTGAATTCAGAGCATGTCAGTGGATAGTAGATCTGTGGCGATATGAG 369
Qy 1895 AGGATGCACTGCTTTCCCAATTCCTGATGAAATGTTGTTATCTAGTGTAACTTT 1954
Db 368 AGGATGCACTGCTTTCCCAATTCCTGATGAAATGTTGTTATCTAGTGTAACTTT 309
Qy 1955 GTAATTTTTTCTAGTGTGTAATGTTGTTGTTTAAATAGTATTATTTTGGCCTTA 2014
Db 308 GTAATTTTTTCTAGTGTGTAATGTTGTTGTTTAAATAGTATTATTTTGGCCTTA 249
Qy 2015 CAATACCGTAACAAATGTTTGTCAATTTTGAATTAATTAATGCAAGTAACTGCTT 2074
Db 248 CAATACCGTAACAAATGTTTGTCAATTTTGAATTAATTAATGCAAGTAACTGCTT 189
Qy 2075 TCGAAATTTGGAAGTGGTGTATTTTATTTTGAAGAAGCAATATGTTTGAATTAATGCTTT 2134
Db 188 TCGAAATTTGGAAGTGGTGTATTTTATTTTGAAGAAGCAATATGTTTGAATTAATGCTTT 129
Qy 2135 GATTGTTTCATATCAAGAAATTCATTGCAAGCTTCTCAACCTGTTTACGGTACTTGGTAA 2194

```

Db 128 GATTGTCATATCAAGAAATGATTGAACGTTCTCAAAACCCGTGTTAGCGTACTGTGTA 69

QY 2195 GAGGAGCCGGTTTGGGAGAACCATTCGCTGCTGCAAGTCTTCTGTTAAAGTCT 2254

Db 68 GAGGAGCCGGTTTGGGAGAACCATTCGCTGCTGCAAGTCTTCTGTTAAAGTCT 9

QY 2255 TTTAACT 2262

Db 8 TTTAACT 1

RESULT 12  
AL577330 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1085YF09 3  
LOCUS AL577330 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1085YF09 3  
DEFINITION prime, mRNA sequence.  
ACCESSION AL577330 GI:12940354  
VERSION AL577330  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 909)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers

FEATURES  
source  
1. 909  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1085YF09"  
/clone\_lib="LTI\_NFL006.PL2"  
/cbase\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime and enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371  
Email: fliang@life.com URL:  
http://fulllength.invitrogen.com;

BASE COUNT 331 a 177 c 136 g 252 t 13 others

Query Match 35.5%; Score 828; DB 9; Length 909;  
Best Local Similarity 95.2%; Pred. No. 1.3e-207;  
Matches 861; Conservative 12; Mismatches 27; Indels 4; Gaps 2;

QY 1394 ACTGCTGGGAACTCTTCCTGGAATCATTCACTGAGATCCCAACCGTGAAGTG 1453

Db 909 ACTGCTGGGAACTCTTCCTGGAATCATTCACTGAGATCCCAACCGTGAAGTG 850

QY 1454 TAACCTGCTCTTTCGCAAAAGATTGATTCGAATTCCTCTGTAAGTTAGCTATAC 1513

Db 849 TAACCTGCTCTTTCGCAAAAGATTGATTCGAATTCCTCTGTAAGTTAGCTATAC 790

QY 1514 AATTGAGAAATAGACGTTCACTGCAACTTTAGTGGTGAGAAATTTAGTTAGT 1573

Db 789 AATTGAGAAATTTASAKTTCACTGCAACTTTAGTGGTGAGAAATTTAGTTAGT 730

QY 1574 GTTGGGATCGACCTCACTTCTGTTGTTCTTTATGTGTGTCTTCTAATCATGAT 1633

Db 729 GTTGGGATCGACCTCACTTCTGTTGTTCTTTATGTGTGTCTTCTAATCATGAT 670

QY 1634 CATAGCCAAAACCTTTTGGAAAGTGTGAGATGTTGTTCTTTAATCCACAGA 1693

Db 669 CATAGCCAAAACCTTTTGGAAAGTGTGAGATGTTGTTCTTTAATCCACAGA 610

QY 1694 AGACATCAAGATACCTGTAATTAAGCTGATAGCATATATTCATACCTGTTAAGCT 1753

Db 609 AGACATCAAGATACCTGTAATTAAGCTGATAGCATATATTCATACCTGTTAAGCT 550

QY 1754 TGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTTATTAACCTGATCTGTAATATGCT 1813

Db 549 TGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTTATTAACCTGATCTGTAATATGCT 490

QY 1814 TGTAGAAAAGCTGTTCCTCATAGCTGTAACAGACTGTAATCAAGATGTAAGTATA 1873

Db 489 TGTAGAAAAGCTGTTCCTCATAGCTGTAACAGACTGTAATCAAGATGTAAGTATA 430

QY 1874 GTAGATCTGTGGCATATAGAGAGATGAGATGCTGCTTCCCATTCATCTGATGAAAT 1933

Db 429 GTAGATCTGTGGCATATAGAGAGATGAGATGCTGCTTCCCATTCATCTGATGAAAT 370

QY 1934 TGTATATCTAGTTAACTTTGTAATTTTCTAGTTGTATGTTAGTCTGTAAT 1993

Db 369 TGTATATCTAGTTAACTTTGTAATTTTCTAGTTGTATGTTAGTCTGTAAT 310

QY 1994 AGGATATATTTTGGCTTACATACCGTAACATGTTGTCATTTTGAATTAAT 2053

Db 309 AGGATATATTTTGGCTTACATACCGTAACATGTTGTCATTTTGAATTAAT 250

QY 2054 GCCAAGTAC--AATGATGCTTGGAAATTTGGAAATGCTTTATCTTGAGAAGA 2111

Db 249 GCCAAGTACATATGATGATGCTTGGAAATTTGGAAATGCTTTATCTTGAGAAGA 190

QY 2112 AATATGTTGATTAATGTTGATGATGTC--ATATCAAGAAATGATGAAGCTTCT 2169

Db 189 AATATGTTGATTAATGTTGATGATGTC--ATATCAAGAAATGATGAAGCTTCT 130

QY 2170 AAACCTGTTTACCGTACTGTTAAGAGGAGCCGGTTTGGAGAGACCTTCATGCT 2229

Db 129 AAACCTGTTTACCGTACTGTTAAGAGGAGCCGGTTTGGAGAGACCTTCATGCT 70

QY 2230 GTCCAGTGTCTTCTGTTAAGTGTCTTTAACTGAGAGGCTAACCTCAAAATCTTTT 2289

Db 69 GTCCAGTGTCTTCTGTTAAGTGTCTTTAACTGAGAGGCTAACCTCAAAATCTTTT 10

QY 2290 TTA 2293

Db 9 CTGA 6

RESULT 13  
AL544755 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1022YJ23 5  
LOCUS AL544755 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1022YJ23 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL544755  
VERSION AL544755 GI:12877235  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 885)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers

FEATURES  
source  
1. 885  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1022YJ23"  
/clone\_lib="LTI\_NFL006.PL2"

/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed by  
 Life Technologies. Contact : Feng Liang Life Technologies,  
 a division of Invitrogen 9800 Medical Center Drive  
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 247 a 211 c 228 g 198 t 1 others  
 ORIGIN

Query Match 35.4%; Score 825; DB 9; Length 885;  
 Best Local Similarity 99.2%; Pred. No. 7.8e-207;  
 Matches 880; Conservative 1; Mismatches 1; Indels 5; Gaps 5;

QY 95 AGGGCTTGAACAGGGCGCTGTAGCGCTTTACCTGTAAAGCAAGATGCAATTTAG 154  
 Db 1 AGGGCTTGAACAGGGCGCTGTAGCGCTTTACCTGTAAAGCAAGATGCAATTTAG 60  
 QY 155 AAATTTTAACTACAGTTTGTAGTCCCTGATTCCTGTGGCAACAGTATCTTCAG 214  
 Db 61 AAATTTTAACTACAGTTTGTAGTCCCTGATTCCTGTGGCAACAGTATCTTCAG 120  
 QY 215 CGAAAGTGAACAGGGCGCTGTAGCGCTTTACCTGTAAAGCAAGATGCAATTTAG 274  
 Db 121 CGAAAGTGAACAGGGCGCTGTAGCGCTTTACCTGTAAAGCAAGATGCAATTTAG 180  
 QY 275 CTTTCAGTATTTAAAGAGCTTCAACAGAGTCAAGTAACTTACGACCCCTTCTCGG 334  
 Db 181 CTTTCAGTATTTAAAGAGCTTCAACAGAGTCAAGTAACTTACGACCCCTTCTCGG 240  
 QY 335 AGCAGATCCAGGCTCCAGCTGCTAAGAGTGTCTTGGGAAAGGAATGAAGTTATA 394  
 Db 241 AGCAGATCCAGGCTCCAGCTGCTAAGAGTGTCTTGGGAAAGGAATGAAGTTATA 300  
 QY 395 TTTTGTCTCAGACCTTACATAGGAGTCAACCTCGCTCCGCAATCAAGCAAGCA 454  
 Db 301 TTTTGTCTCAGACCTTACATAGGAGTCAACCTCGCTCCGCAATCAAGCAAGCA 360  
 QY 455 GTTTCTGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 514  
 Db 361 GTTTCTGATCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 420  
 QY 515 CCCAGTCAAACTATGATCTCTATATGCTATCTCAAGCTGGGCGGCAAGGAAAGTA 574  
 Db 421 CCCAGTCAAACTATGATCTCTATATGCTATCTCAAGCTGGGCGGCAAGGAAAGTA 480  
 QY 575 TGAATTCAGCGAGTCAACACCACTCCAGCGGTGGTCCATGTATGTAGAGTGA 634  
 Db 481 TGAATTCAGCGAGTCAACACCACTCCAGCGGTGGTCCATGTATGTAGAGTGA 540  
 QY 635 TCAAGAGAGGAGGAG 693  
 Db 541 TCAAGAGAGGAGGAG 600  
 QY 694 TCCAGACAGGAGGCGGAGTACACCCGATCCACCTCAGCTGMACTGGCCAGCAGAG 753  
 Db 601 TCCAGACAGGAGGCGGAGTACACCCGATCCACCTCAGCTGMACTGGCCAGCAGAG 660  
 QY 754 GACGCAATCCAAATCATCTACGGAGAGATCTTTTACTGTGGAGGTGGTGGTCA 813  
 Db 661 GACGCAATCCAAATCATCTACGGAGAGATCTTTTACTGTGGAGGTGGTGGTCA 720  
 QY 814 CTTCTTCGAGGTGGCAGCGGAGTCCGGGTGGCAGAGATCCAGTTCATGTGCTCAGA 873  
 Db 721 CTTCTTCGAGGTGGCAGCGGAGTCCGGGTGGCAGAGATCCAGTTCATGTGCTCAGA 780  
 QY 874 AGGAATCAAGGCGGTGTCCTTCTTAATCTGTCACACAGTACTGTTTCATGGCAG 933  
 Db 781 AGGAATCAAGGCGGTGTCCTTCTTAATCTGTCACACAGTACTGTTTCATGGCAG 838

QY 934 CCGGAATGACTTGGGCAATCACTAGTTGTGTGATCGCACAG 980  
 Db 839 CCGGAATGACTTGGG-CAATCACTAG-TTGTGTGATCGCACAG 883

RESULT 14  
 AL566338/c  
 LOCUS  
 DEFINITION  
 AL566338 LTI\_FL013\_FBrnl 849 bp mRNA linear EST 16-FEB-2001  
 prime, mRNA sequence.  
 ACCESSION  
 AL566338  
 VERSION  
 AL566338.1 GI:12918597  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 849)  
 AUTHORS  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE  
 Full-length cDNA libraries and normalization  
 JOURNAL  
 Unpublished (2001)  
 COMMENT  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: sef@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1. 849  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CSODF015F21"  
 /clone\_lib="LTI\_FL013\_FBrnl"  
 /dev\_stage="pooled tissue from post conception fetuses (20  
 week, 24 week and 26 week)"  
 /lab\_host="DH10B"  
 /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand  
 cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-stranded cDNA was digested with Not I  
 and cloned into the Not I and Eco RV sites of the  
 pCMVSPORT 6 vector. Library was constructed by Life  
 Technologies. Contact : Feng Liang Life Technologies, a  
 division of Invitrogen 9800 Medical Center Drive, Rockville  
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
 fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 309 a 175 c 120 g 235 t 10 others  
 ORIGIN

Query Match 34.7%; Score 808.6; DB 9; Length 849;  
 Best Local Similarity 97.5%; Pred. No. 1.7e-202;  
 Matches 830; Conservative 10; Mismatches 9; Indels 2; Gaps 2;

QY 1450 CGTCTACTTGTGCTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTT 1509  
 Db 849 CGTGTACTTGTGCTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTT 790  
 QY 1510 ATACAATTCAGAAATAGCAGTTTCACTGCCACTTTTGTAGTGGTGAGAAATTTAGTTT 1569  
 Db 789 ATACAATTCAGAAATAGCAGTTTCACTGCCACTTTTGTAGTGGTGAGAAATTTAGTTT 731  
 QY 1570 AGTGTGTGGATCCGACCTTCTGTGTTCTTTATGTGGTGGTCTATACAT 1629  
 Db 730 AGTGTGTGGATCCGACCTTCTGTGTTCTTTATGTGGTGGTCTATACAT 671  
 QY 1630 GAATCATAGCCAAAATTTTGGAACTGTTGGTGGAGATGTTGTTTACCC 1689  
 Db 670 GAATCATAGCCAAAATTTTGGAACTGTTGGTGGAGATGTTGTTTACCC 611  
 QY 1690 ACCAAGACATCAAGATACACTTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGA 1749  
 Db 610 ACCAAGACATCAAGATACACTTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGA 551  
 QY 1750 CACTTGGGTGAAAGATATGGCAGTGGGAGACTAAGATGTATTAACCTACCTGTAATCAT 1809

```

|||||
Db 550 CACTGGTGAAGATATGCGAGTGGAGACATGATATTAACCTACCTGTAACAT 491
QY 1810 ATGTTGTGGAAGACCTTCCCATGCTAAACAGACTGTAATCAAGACATGCAAGT 1869
Db 490 ATGTTGTGGAAGACCTTCCCATGCTAAACAGACTGTAATCAAGACATGCAAGT 431
QY 1870 GATAGTAGTCTGTGGCGATATAGAGGAGATGACAGTGCCTTCCCATCTATCTGATG 1929
Db 430 GATAGTAGTCTGTGGCGATATAGAGGAGATGACAGTGCCTTCCCATCTATCTGATG 371
QY 1930 GAATGTTATCTAGTGAATTAACCTTGAATTTTCTAGTGAATGATGATGCTGCT 1989
Db 370 GAATGTTATCTAGTGAATTAACCTTGAATTTTCTAGTGAATGATGATGCTGCT 311
QY 1990 AATAGGTATTAATTTTGGCCTTACATACCGTACATGTTTGTCAATTTGAAATCT 2049
Db 310 AATAGGTATTAATTTTGGCCTTACATACCGTACATGTTTGTCAATTTGATATCT 251
QY 2050 TAATGCCAAGTAAACAATGATGCTTGGAAATTTGGAAATGTTTATCTTTAGAG 2109
Db 250 TAATGCCAAGTAAACAATGATGCTTGGAAATTTGGAAATGTTTATCTTTAGAG 191
QY 2110 CAATATGTTTGAATTAATGCTTGTGATGCTATCAATCAAGAAATGATTGAACGTTCTC 2169
Db 190 CAATATGTTTGAATTAATGCTTGTGATGCTATCAATCAAGAAATGATTGAACGTTCTC 131
QY 2170 AAACCTGTTTACGCTACTTGTGAAGAGGAGCCGCTTTGGAGAGACATGTCATCGCT 2229
Db 130 AAACCTGTTTACGCTACTTGTGAAGAGGAGCCGCTTTGGAGAGACATGTCATCGCT 72
QY 2230 GTCCAACTGTTTCTTCTTGAATGCTTTTAACTGGAGAGGCTAACCTCAAAATATCTTTT 2289
Db 71 GTCCAACTGTTTCTTCTTGAATGCTTTTAACTGGAGAGGCTAACCTCAAAATATCTTTT 12
QY 2290 TTAACTGCATT 2300
Db 11 TTAACTGCATT 1

RESULT 15
BG695987 856 bp mRNA linear EST 07-MAY-2001
LOCUS BG695987
DEFINITION BG695987 Homo sapiens cDNA clone IMAGE:480918 5',
mRNA sequence.
ACCESSION BG695987
VERSION BG695987.1 GI:13960662
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsrbs-remail.nih.gov.
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jhmi.gov
Plate: LLM10692 row: 1 column: 07
High quality sequence stop: 821.
Location/Qualifiers
1..856
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:480918"
/clone_1ib="NCI_CGAP_Skin3"

```

```

/lab host="DH10B (T1 phage-resistant)"
/Note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 230 a 148 c 174 g 304 t
ORIGIN
Query Match 34.6%; Score 806.6; DB 12; Length 856;
Best Local Similarity 98.6%; Pred. No. 5.7e-202;
Matches 845; Conservative 0; Mismatches 9; Indels 3; Gaps 3;
QY 1208 GTAGCGTTCTCTTCCCTCATGCTTCTGATGCTTGGCATATATCTGATTTACA 1267
Db 1 GTAGCGTTCTCTTCCCTCATGCTTCTGATGCTTGGCATATATCTGATTTACA 60
QY 1268 AGACTAACCTTGTGTATATTAAGTACACCGTGTGTTTACATCTTTGGAGAC 1327
Db 61 AGACTAACCTTGTGTATATTAAGTACACCGTGTGTTTACATCTTTGGAGAC 120
QY 1328 CAGGAAAGCGTTGGAAACGATACCTTTCCGACATTCGGATCTTCCGACTTTTGC 1387
Db 121 CAGGAAAGCGTTGGAAACGATACCTTTCCGACATTCGGATCTTCCGACTTTTGC 180
QY 1388 AACGCACTTGTCTGGGAACTCTCTGGAAATGCAATCAGAGATCCCAACCGTGC 1447
Db 181 AACGCACTTGTCTGGGAACTCTCTGGAAATGCAATCAGAGATCCCAACCGTGC 240
QY 1448 AACGCTAACCTTGTCTTGGCAAAAGAGATGATGAAATTCCTCTGTGAATTTAGC 1507
Db 241 AACGCTAACCTTGTCTTGGCAAAAGAGATGATGAAATTCCTCTGTGAATTTAGC 300
QY 1508 TTATCAATTCAGAGATAGACATTTTCACTGCCAATCTTTAGGGGAGAAATTTAGT 1567
Db 301 TTATCAATTCAGAGATAGACATTTTCACTGCCAATCTTTAGGGGAGAAATTTAGT 360
QY 1568 TTAGGTGTTGGAGTCGGACCTCAGTTCTGTGTTTCTTTATGAGTGAGTTCTATAC 1627
Db 361 TTAGGTGTTGGAGTCGGACCTCAGTTCTGTGTTTCTTTATGAGTGAGTTCTATAC 420
QY 1628 ATGAATCATAGCCAAACCTTTTGGAACTGTTGTTGAGATAGTTGTTCTTTTACC 1687
Db 421 ATGAATCATAGCCAAACCTTTTGGAACTGTTGTTGAGATAGTTGTTCTTTTACC 480
QY 1688 CCACGAAAGATCAAGATACCTGTTAAATGAAGCTATGATATTCATACCTGTTG 1747
Db 481 CCACGAAAGATCAAGATACCTGTTAAATGAAGCTATGATATTCATACCTGTTG 540
QY 1748 TACACTTGGGTGAAAGATATGAGCAGTGGAGACTAAGATGATTNACTTACTGTAATC 1807
Db 541 TACACTTGGGTGAAAGATATGAGCAGTGGAGACTAAGATGATTNACTTACTGTAATC 600
QY 1808 ATATGTTGTGAGAAA-GCTGTTCCCATGCTTCAAGAGACTTGAATCAAGACATGCA 1866
Db 601 ATATGTTGTGAGAAA-GCTGTTCCCATGCTTCAAGAGACTTGAATCAAGACATGCA 660
QY 1867 GTGAGATAGATCTGTGGCATATGAGAGGAGATCAGTGCCTTCCCATCTTCTCTG 1926
Db 661 GTGAGATAGATCTGTGGCATATGAGAGGAGATCAGTGCCTTCCCATCTTCTCTG 720
QY 1927 ATGAAATGTTATAGTAGTTTACA-TTGTATATTTTCTTCTAGTTGTAATGTATGTC 1985
Db 721 ATGAAATGTTATAGTAGTTTACA-TTGTATATTTTCTTCTAGTTGTAATGTATGTC 780
QY 1986 TGTGAATAGTATATATTTTGGCCTTCAATACCGTAACATGTTTGTCAATTTGAAA 2045
Db 781 TGTGAATAGTATATATTTTGGCCTTCAATATAGTAAACATG-TTGTCAATTTGAAA 839
QY 2046 TACTTAATGCCAAGTAA 2062
Db 840 TACTTAATGCCAAGTAA 856

```

Sun Dec 15 08:38:04 2002

Search completed: December 14, 2002, 18:11:18  
Job time : 2612.11 secs

us-09-782-953-11.rst

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 97.497 Seconds  
(without alignments)  
13789.580 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggatttttaggacttag.....cacogatccaccttagctga 597

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq 101002:\*

- 1: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneqseq-emb1/NA1988.DAT:\*
- 10: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1989.DAT:\*
- 11: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1990.DAT:\*
- 12: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1991.DAT:\*
- 13: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1992.DAT:\*
- 14: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1993.DAT:\*
- 15: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1994.DAT:\*
- 16: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1995.DAT:\*
- 17: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1996.DAT:\*
- 18: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1997.DAT:\*
- 19: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1998.DAT:\*
- 20: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA1999.DAT:\*
- 21: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA2000.DAT:\*
- 22: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA2001A.DAT:\*
- 23: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA2001B.DAT:\*
- 24: /SID22/gcgdata/geneqseq/geneqseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	100.0	597	24	AAD30153
2	515.8	86.4	597	24	AAD30152
3	504.4	84.5	599	24	AAD30151
4	470.6	78.8	2331	24	AAD30155
5	459.6	77.0	2348	24	AAS94805
6	408.2	68.4	2212	24	AAD30156
7	399.6	66.9	2173	24	ABN96843
8	398	66.7	2174	20	AAK01282
9	395.4	66.2	2358	21	AAF18328

10	273.8	45.9	442	21	AAC01774	Human secreted pro
11	214.2	35.9	3184	24	AAD30157	Human MCIP associa
12	214.2	35.9	3184	24	ABL61768	Colon adenocarcino
13	214.2	35.9	3184	24	ABL65187	Lung cancer relate
14	209.6	35.1	934	22	AAF25338	Nucleotide sequenc
15	209.6	35.1	3159	24	ABA91457	Human Down syndrom
16	201.4	33.7	594	24	AAD30154	Mouse MCIP associa
17	191.2	32.0	720	24	AAD30159	Human MCIP associa
18	191.2	32.0	828	24	AAD30158	Human MCIP associa
19	186.2	31.2	615	24	ABA91463	Rat Down syndrome
20	155	26.0	1021	24	ABA91463	Rat Down syndrome
21	148	24.8	840	24	ABA91458	Human colon cancer
22	144.6	24.2	412	22	ABA56172	Human foetal liver
23	144.6	24.2	412	22	AAK04364	Human brain expres
24	144.6	24.2	412	22	AAI14449	Probe #4382 for ge
25	144.6	24.2	412	22	AAI135821	Probe #4507 used t
26	144.6	24.2	412	22	AAI04272	Probe #4263 used t
27	144.6	24.2	412	22	ABSO4419	Human genome-deriv
28	144.6	24.2	446	22	ABA43406	Human breast cell
29	144.6	24.2	446	22	ABA53852	Human foetal liver
30	144.6	24.2	446	22	ABA53852	Probe #2064 for ge
31	144.6	24.2	446	22	AAK02113	Human brain expres
32	144.6	24.2	446	22	AAK27563	Human bone marrow
33	144.6	24.2	446	22	AAI12148	Probe #2081 for ge
34	144.6	24.2	446	22	AAI33493	Probe #2179 used t
35	144.6	24.2	446	22	AAI02061	Probe #2052 used t
36	144.6	24.2	446	24	ABSO2041	Human genome-deriv
37	123	20.6	486	22	ABA42132	Human breast cell
38	123	20.6	486	22	ABA52554	Human foetal liver
39	123	20.6	486	22	ABA22343	Probe #809 for gen
40	123	20.6	486	22	AAK00816	Human brain expres
41	123	20.6	486	22	AAK26270	Human bone marrow
42	123	20.6	486	22	AAI10903	Probe #836 for gen
43	123	20.6	486	22	AAI32163	Probe #849 used to
44	123	20.6	486	22	AAI00825	Probe #816 used to
45	123	20.6	486	24	ABSO0858	Human genome-deriv

## ALIGNMENTS

## RESULT 1

AAD30153  
ID AAD30153 standard; DNA; 597 BP.

XX AAD30153;

XX AC AAD30153;

XX DT 17-MAY-2002 (first entry)

XX DE Mouse MCIP associated DNA #3.

XX OS Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

XX KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.

XX OS Mus musculus.

XX FT Key Location/Qualifiers

XX FT CDS 1..597

XX FT FT /\*tag= a

XX FT /product= "Mouse MCIP associated protein #3"

XX PN WO200204491-A2.

XX PD 17-JAN-2002.

XX PF 06-JUL-2001; 2001WO-US21662.

XX PR 07-JUL-2000; 2000US-216601P.

XX PR 13-FEB-2001; 2001US-0782953.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PA (WILL) WILLIAMS S R.

XX PA (ROTH) ROTHERMEL B.

XX Williams SR, Rothermel B;  
 XX WPI; 2002-179698/23.  
 DR P-PSDB; AAE18912.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 XX Disclosure: Page 150-151; 174pp; English.  
 XX  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated DNA.  
 CC Note: This sequence has been described as human MCIP encoding DNA  
 CC in the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.  
 XX  
 XX Sequence 597 BP; 170 A; 156 C; 142 G; 129 T; 0 other;  
 XX  
 XX Query Match 100.0%; Score 597; DB 24; Length 597;  
 XX Best Local Similarity 100.0%; Pred. No. 5,3e-176;  
 XX Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTTTATGGGACTTAGCTACATTTTATGCTCCTGATTCCTGTTGGCAACGAT 60  
 DB 1 ATGATTTTATGGGACTTAGCTACATTTTATGCTCCTGATTCCTGTTGGCAACGAT 60  
 QY 61 GATGCTTCACGGAAGTAGACAGGAGCCCAATTTGATTCCTCTTCAGAAATATGAC 120  
 DB 61 GATGCTTCACGGAAGTAGACAGGAGCCCAATTTGATTCCTCTTCAGAAATATGAC 120  
 QY 121 AAGGACACCACTTCGATATTTTAAAGGCTTCAACGTCGCGATTAACCTTCAGCAAC 180  
 DB 121 AAGGACACCACTTCGATATTTTAAAGGCTTCAACGTCGCGATTAACCTTCAGCAAC 180  
 QY 181 CCCTATCTGACGCGATGCGAGGCTGCGCTGCACAAAGCCGAGTTCCTGGGAAAGAA 240  
 DB 181 CCCTATCTGACGCGATGCGAGGCTGCGCTGCACAAAGCCGAGTTCCTGGGAAAGAA 240  
 QY 241 ATGAAGTTGATTTTGTCTCAGCTTATACATAGGAAGTTTACACCTGCTCCGCCAAT 300  
 DB 241 ATGAAGTTGATTTTGTCTCAGCTTATACATAGGAAGTTTACACCTGCTCCGCCAAT 300  
 QY 301 CCCGCAAAACAGTCTCATCTCCCTCCGCGCTCTCTCCGTTGCTGGAACAGTA 360  
 DB 301 CCCGCAAAACAGTCTCATCTCCCTCCGCGCTCTCTCCGTTGCTGGAACAGTA 360  
 QY 361 GAAGATGCCACCCCGCTCATTAATTCATTTTATGATTCATCCAGCTGGGGCCA 420  
 DB 361 GAAGATGCCACCCCGCTCATTAATTCATTTTATGATTCATCCAGCTGGGGCCA 420  
 QY 421 GGAAGAAGATTAAGTATGATGATGAGGAGACCCCACTCCAGTGTGTGCTCAGCTG 480  
 DB 421 GGAAGAAGATTAAGTATGATGATGAGGAGACCCCACTCCAGTGTGTGCTCAGCTG 480  
 QY 481 TGTGAAGTACCAAGAGAAATGAGAGAAAGAGAAAGATGAGAAATGAAGAGACCC 540  
 DB 481 TGTGAAGTACCAAGAGAAATGAGAGAAAGAGAAAGATGAGAAATGAAGAGACCC 540  
 QY 541 AAGCCCAAAATCATCGAGACAGGAGACGGAGTACACACCGATCCACCTTAGCTGA 597

DB 541 AAGCCCAAAATCATCGAGACAGGAGACGGAGTACACACCGATCCACCTTAGCTGA 597  
 XX  
 XX RESULT 2  
 XX AAD30152  
 XX ID AAD30152 standard; DNA; 597 BP.  
 XX  
 XX AAD30152;  
 XX  
 XX 17-MAY-2002 (first entry)  
 XX  
 XX Human MCIP associated DNA #2.  
 XX  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 1..597  
 XX FT /\*tag= a  
 XX FT /product= "Human MCIP associated protein #2"  
 XX  
 XX MO200204491-A2.  
 XX  
 XX 17-JUN-2002.  
 XX  
 XX 06-JUL-2001; 2001WO-US21662.  
 XX  
 XX 07-JUL-2000; 2000US-216601P.  
 XX  
 XX 13-FEB-2001; 2001US-0782953.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 XX (WILL/) WILLIAMS S R.  
 XX (ROTH/) ROTHERMEL B.  
 XX  
 XX Williams SR, Rothermel B;  
 XX WPI; 2002-179698/23.  
 XX P-PSDB; AAE18911.  
 XX  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 XX Claim 95; Page 147-148; 174pp; English.  
 XX  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 XX  
 XX Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;  
 XX  
 XX Query Match 86.4%; Score 515.8; DB 24; Length 597;  
 XX Best Local Similarity 94.4%; Pred. No. 1.2e-150;  
 XX Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 31 AGCTCCCTGATTCCTGTGTGCGCAACGATGATGCTTCAGCGAAGTAGACAGGCGC 90  
 DB 31 AGCGCCACCATTCCTGCGCACCTGACCCCGCGGTGTTGTGAGAGGCGCTGTGCGGCGC 90  
 QY 91 AAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCAGTATTTTAAAGAC 150



Db 91 "AAATTGAATCCCTCTCAGAACATATGACAAAGGACACACCTTCCAGTATTTAAAGAGC 150  
 Qy 151 TTCAAACGTGTCGGGATAACTTACGAAACCCCTTATCTGCAGCGATGCCAGGCTGGG 210  
 Db 151 TTCAAACGTGTCGGGATAACTTACGAAACCCCTTATCTGCAGCGATGCCAGGCTGGG 210  
 Qy 211 CTGCAACAGCCGAGTTCCTGGGAGGAAATGAAGTTGTTTGTCTCAGACTTTACAC 270  
 Db 211 CTGCAACAGCCGAGTTCCTGGGAGGAAATGAAGTTGTTTGTCTCAGACTTTACAC 270  
 Qy 271 ATAGAAGTTTACACCTCGCTCCGCCCAATCCGACAAACAGTTTCTCATCTCCCTCCG 330  
 Db 271 ATAGAAGTTTACACCTCGCTCCGCCCAATCCGACAAACAGTTTCTCATCTCCCTCCG 330  
 Qy 331 GCCTCTCCTCCGTTGGTGGAAACAAGTAGAAGATGCACCCCGCTCATAAATTCAGAT 390  
 Db 331 GCCTCTCCTCCGTTGGTGGAAACAAGTAGAAGATGCACCCCGCTCATAAATTCAGAT 390  
 Qy 391 CTTTATATGCAATCTCAAGCTGGGCGCAGGAGAGATGATGAATCTGCATGCGAGCA 450  
 Db 391 CTTTATATGCAATCTCAAGCTGGGCGCAGGAGAGATGATGAATCTGCATGCGAGCA 450  
 Qy 451 GACCCCACTCCCACTGGTGGTCCAGTGTGTCAGAGTGACCAAGAGAAATGAGGAGAA 510  
 Db 451 GACCCCACTCCCACTGGTGGTCCAGTGTGTCAGAGTGACCAAGAGAAATGAGGAGAA 510  
 Qy 511 GAGGAAGATGAGAGAGATGAAGAGACCCCAAGCCCAAAATCATCCAGACGAGACCG 570  
 Db 511 GAGGAAGATGAGAGAGATGAAGAGACCCCAAGCCCAAAATCATCCAGACGAGACCG 570  
 Qy 571 GAGTACACACCGATCCACCTTAGCTGA 597  
 Db 571 GAGTACACACCGATCCACCTTAGCTGA 597

## RESULT 3

AAD30151  
 ID AAD30151 standard; DNA; 599 BP.  
 XX  
 AC AAD30151;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human MCIP associated DNA #1.  
 XX  
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204491-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 06-JUL-2001; 2001WO-US21662.  
 XX  
 PR 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX  
 PI Williams SR, Rothermel B;  
 XX  
 XX WPI; 2002-179698/23.  
 XX  
 PT Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 147; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.

SQ Sequence 599 BP; 164 A; 150 C; 149 G; 136 T; 0 other;

Query Match 84.5%; Score 504.4; DB 24; Length 599;

Best Local Similarity 99.6%; Pred. No. 4.3e-147;

Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 4 GATTTTAGGAGCTTTAGCTAGCAATTTAGCTCCCTGATTTGTTGGCAACCATGAT 63

Db 83 GATTTTAGGAGCTTTAGCTAGCAATTTAGCTCCCTGATTTGTTGGCAACCATGAT 142

Qy 64 GTCTTCAGCGAAAGTGAGACAGGCGCAAAATTTGAATCCCTTTCAGACATATCACAAG 123

Db 143 GTCTTCAGCGAAAGTGAGACAGGCGCAAAATTTGAATCCCTTTCAGACATATCACAAG 202

Qy 124 GACACACCTTCCAGTATTTTAAGAGCTTCAAACTGTCGGGATAAACTTCAGCAACCCC 183

Db 203 GACACACCTTCCAGTATTTTAAGAGCTTCAAACTGTCGGGATAAACTTCAGCAACCCC 262

Qy 184 TTATCTGCAGCGGATGCCAGGCTGCGGTGCACAAGACCGAGTTCTGGGGAAGAAATG 243

Db 263 TTATCTGCAGCGGATGCCAGGCTGCGGTGCACAAGACCGAGTTCTGGGGAAGAAATG 322

Qy 244 AAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAATCCC 303

Db 323 AAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAATCCC 381

Qy 304 GACAAACAGTTCTCATCTCCCTCCCGCTCTCTCCCGTTGGCTGGAAACAAGTAGAA 363

Db 382 GACAAACAGTTCTCATCTCCCTCCCGCTCTCTCCCGTTGGCTGGAAACAAGTAGAA 441

Qy 364 GATGCCACCCCGTCATAAATTACGATCTTTTATATGCCATCTCAAGCTGGGCGCAGGA 423

Db 442 GATGCCACCCCGTCATAAATTACGATCTTTTATATGCCATCTCAAGCTGGGCGCAGGA 501

Qy 424 GAGAAAGTATGAAGTGCATGCGAGCGACAGACCCCACTCCAGTGTGGTCCACGTGTGT 483

Db 502 GAGAAAGTATGAAGTGCATGCGAGCGACAGACCCCACTCCAGTGTGGTCCACGTGTGT 561

Qy 484 GAGAGTGACCAAGAGATGAGGAGGAAGAGAGAT 521

Db 562 GAGAGTGACCAAGAGATGAGGAGGAAGAGAGAT 599

## RESULT 4

AAD30155  
 ID AAD30155 standard; DNA; 2331 BP.

XX AAD30155;

XX AC

XX 17-MAY-2002 (first entry)

XX Human MCIP associated DNA #1.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX  
 OS Homo sapiens.

XX Key  
 FH Location/Qualifiers

Db 444 CCAGCAAGCAGGTTCTGATCTCCCTCCGCGCCTCTCCGCGAGGAGGATGAAACAATG 503

0y 361 GAAAGTGCACACCCCGTCATAAATTAAGATCTTTTATATGCAATCTCCAACTGGAGCCA 420

Db 504 GAAAGTGGACCCCGATCATTAACATATGATCTTTATATGCAATCTCCAACTGGAGCCA 563

0y 421 GGAGAGAGATATGAACCTGATGATGAGCGGACGACACCCCACTCCAGTGGTGTCCAGCTG 480

Db 564 GCGGAAAGATATGAATTTGACACGACCGCATGTACACCACTCCAGGTGTGTGTCCATGTA 623

0y 481 TGTGAGAGTGCACCAAGAGATGAGAGGAAAGAGAAAGATGAGAGAAATGAAGAGACC 540

Db 624 TGTGAGAGTGCATCAAGAGAGAGAGAAAGAGAA--ATGAGAAAGATGAGAGAGACTT 680

0y 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTGAAGCTGA 597

Db 681 AAGCCCAAAATTTATTCACACCGAGAGGAGGAGTACACCGATCCACTTGAAGCTGA 737

RESULT 5  
AAS94805  
ID AAS94805 standard; DNA; 2348 BP.

XX AAS94805;  
XX  
XX  
DT 14-FEB-2002 (first entry)

XX Human DNA sequence #60 expressed during foam cell differentiation.  
XX  
DE Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
XX  
XX Human; foam cell differentiation; coronary artery disease; gene therapy; ds.  
XX  
XX  
XX  
XX Homo sapiens.  
OS  
PN WO200177389-A2.  
XX  
XX 18-OCT-2001.  
PD  
XX  
XX 04-APR-2001; 2001WO-US11128.  
PF  
XX  
XX 05-APR-2000; 2000US-195106P.  
PR  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
PI Shiffman D, Somogyi R, Lawn R, Sellhammer JJ, Porter GJ, Mikita T;  
PI Tai J;  
XX  
XX  
XX WPI; 2002-010925/01.  
DR  
XX  
XX  
XX Composition useful for diagnosis of conditions, disorders or diseases  
PT associated with atherosclerosis, comprises several polynucleotides that  
PT are differentially expressed in foam cell development -  
PT  
XX  
XX  
XX Claim 1; Page 112-113; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide  
CC sequences that are differentially expressed during foam cell  
CC differentiation. The polynucleotide sequences of the invention or a  
CC composition comprising these polynucleotides are useful as a high  
CC throughput method for detecting altered expression of one or more  
CC polynucleotides in a sample. The polynucleotides can be used in the  
CC diagnosis of disorders associated with foam cell development such as  
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
CC coronary artery disease. The polynucleotide sequences of the invention are also  
CC useful in gene therapy. AAS94746-AAS95021 represent the human  
CC polynucleotide sequences of the invention which are differentially  
CC expressed during foam cell differentiation.  
XX  
XX  
XX Sequence 2348 BP; 638 A; 473 C; 550 G; 687 T; 0 other;

Query Match 77.0%; Score 459.6; DB 24; Length 2348;  
Best Local Similarity 87.8%; Pstd No. 8.4e-133;





CC q22.1-22.2 region of human chromosome 21. An increase in the transient  
 CC expression of DSCR1 mRNA in the brains of young rats, compared to  
 CC expression levels in the brains of adult rats, suggests an important role  
 CC for DSCR1 during the development of the Central Nervous System (CNS), and  
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities  
 CC of mental retardation and/or heart defects as found in Down's syndrome  
 CC patients.  
 XX

SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;

Query Match 66.7%; Score 398; DB 20; Length 2174;  
 Best Local Similarity 87.6%; Pred. No. 1.4e-113;  
 Matches 447; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

OY 88 GCCAAATTGAATCCCTCTTGGAACATATGACAAAGGACACACCTTCCAGTATTTTAAAG 147  
 DB 58 GCCAAATTGAATCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTCAGTATTTTAAAG 117  
 OY 148 AGCTTCAAAACGTGTCGGATAACTTCAGCAACCCCTTATCTGCAGCGATGCCAGGCTG 207  
 DB 118 AGCTTCAAAACGTGACATTAACCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177  
 OY 208 CGGTGCAACAGACCGAGTTCCTGGGGAAGGAATGAAGTTGTATTTTGTCTCAGACTTTA 267  
 DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGGAATGAAGTTATATTTTGTCTCAGACCTTA 237  
 OY 268 CACATGGAAGTTACACCTGGCTCGGCCAAATCCCGCAACAGTTCCTCATCTCCCT 327  
 DB 238 CACATGGAAGCTCACACCTGGCTCGGCCAAATCCAGCAAGCAGTTTCTGATCTCCCT 297  
 OY 328 CCGGCTCTCTCCGTTGGTGGCAACAGTAGAAGATGCCACCCCGCTCATTAATAC 387  
 DB 298 CCGGCTCTCGCAGTGGATGGAACAACCTGGAAGATGCGACCCAGTCATAAATAT 357  
 OY 388 GATCTTTATATGCATCTCCAAGCTGGGCGCAGGAGAGATGATGAATGCATGCAAGCG 447  
 DB 358 GATCTTTATATGCATCTCCAAGCTGGGCGCAGGAGAGATGATGAATGCATGCAAGCG 417  
 OY 448 ACAGACCCCACTCCAGTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGATGAGGAG 507  
 DB 418 ACTGACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAA 477  
 OY 508 GAAGAGGAGAGATGGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACAGGGA 567  
 DB 478 GAAGAGGAA---ATGGGAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACAGGAGG 534  
 OY 568 CCGAGTACACCGATCCACCTTACCTGA 597  
 DB 535 CCGAGTACACCGATCCACCTCAGCTGA 564

## RESULT 9

AAF18328  
 ID AAF18328 standard; DNA; 2358 BP.

XX AAF18328;

XX 14-MAR-2001 (first entry)

DE Lung cancer associated polynucleotide sequence SEQ ID 347.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease; ds.

OS Homo sapiens.

XX WO200055180-A2.

XX 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

PA Ruben SM;

XX WPI; 2000-587514/55.

DR P-PSDB; AAB58452.

XX Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as lung cancer.

XX Claim 1; Page 805-806; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer

CC associated proteins and polynucleotide sequences, their agonists, and

CC antagonists may have neuroprotective, cytostatic, cardioactive;

CC immunomodulatory; muscular active general; vulnary; gastrointestinal

CC general; nephrotropic; antiinfective; gynecological; or antibacterial

CC activity. The invention also includes antibodies specific for the

CC protein or polynucleotide sequences. The lung cancer associated

CC polynucleotide sequences may be used for detection of lung cancer,

CC chromosome identification, as chromosome markers, and for numerous other

CC diagnostic or research purposes. The proteins may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders. The proteins may also be used in the treatment of wounds and

CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and

CC peptide AAB58549 are used in the course of the invention for the

CC identification and characterization of the polynucleotide and protein

XX sequences.

SQ Sequence 2350 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

## Query Match

Best Local Similarity 66.2%; Score 395.4; DB 21; Length 2358;

Matches 483; Conservative 2; Mismatches 108; Indels 4; Gaps 2;

OY 1 ATGGATTTTAGGGACTTTAGCTACAAATTTTAGCTCCCTGATTGTTGTGTCGCAACCAT 60

DB 114 ATGGAGGAGGTGGACTCGAGGACCTGCCAGCCACCATCGCTGTACCTGACCTGACCGG 173

OY 61 GATGTCTTCAGCGAAAGTGAGACAGGCGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120

DB 174 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 233

OY 121 AAGGACACCACTTCCAGTATTTTAAGAGCTTCAACGTGTCCGGATAACTTCAGCAAC 180

DB 234 AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAACAGGTGAGATTAATTCAGCAAC 293

OY 181 CCCTTATTCAGCGCATGCCAGGCTGCGGCTGCACAGACCGAGTTCCTCGGGAAGAA 240

DB 294 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTCGGGAAGAA 353

OY 241 ATGAAGTTGATTTTGTCTCAGCTTTTACATAGGAAGTTTCAACCTGGCTCCGCCCAAT 300

DB 354 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGTTTCAACCTGGCTCCGCCCAAT 412

OY 301 CCCGACAAACAGTTTCCCTCATCTCCCTCCGCTCTCTCCCGTTGGCTGGAACAAGTA 360

DB 413 CCAGACAGCAGTTTCTGATCTCCCTCCGCTCTCCGAGGAGTGGAAACAAGTG 472

OY 361 GAAGATGCAACCCCGTCATAAATTTAGCATCTTTTATATGCATCTCAAGCTGGGCCA 420

DB 473 GAAGATGCGACCCAGTCATAAATTTAGTCTTTATATGCCATCTCAAGCTGGGCCA 532

OY 421 GGAGAGAGTATGAATGCTGATGCGAGCAGACGCCCACTCCAGTGTGGTTCACGTG 480







CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Query Match 35.9%; Score 214.2; DB 24; Length 3184;

Best Local Similarity 66.1%; Pred. No. 4.1e-56;

Matches 331; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

OY 21 CTACAAATTTAGCTCCCGATGCTGTGTGGCAACATGATGCTTCTTACAGGAAATGA 80  
 DB 210 CTGTGATGTTCTCAGCTTGTGTGCTGTGTGTGTGATGTCAGAGCTTCTTACCAATCAGAGA 269  
 OY 81 GACACGAGGCAATTTGAAATCCCTCTTCAAGACATATGACAGACACACCTTCCAGTA 140  
 DB 270 GGTTAAGGAAAAATTTGGGGGAGCTGTTTCGACTTATGATGACTGTGACGTTCCAGCT 329  
 OY 141 TTTTAAGAGCTTCAAAAGCTGTGCGATTAATCTTCAAGCAACCCCTTATTTGACGCGAGTCC 200  
 DB 330 ATTTAAGAGTTTCAGACGTCGTCGTAATTAATCTTCAAGCAATCTTAAATCTGACGCCGAGC 389  
 OY 201 CAGGCTGCGGCTGCACAGACCGAGTCTCGGAGAGAAATGAAGTTGATTTTGTCTCA 260  
 DB 390 TAGGATAGAGCTTCAATGAACCCCAATTCAGAGGAGAAAAATTAAGCTTCTACTTTGCACA 449  
 OY 261 GACT-----TTACATATGAAAGTTTCAACCTGAGCTCCGCCCAATCCGACAA 308  
 DB 450 GGTTCAGACTCCAGAGACAGATGAGAGACAAACCTGCTCCACCCGACCTGCCA 509  
 OY 309 ACAGTCTCTCATCTCCCTCCGCTCTCTCTCCGTTGGCTGGAACAAAGTGAAGATGC 368  
 DB 510 ACAGTCTCTCATCTCCCTCCCTCTCTCTCCGCTGTTACTGCGAGCCCATCAAGATGC 569  
 OY 369 CACCCCGCTCAATAATTAAGATCTTTATATGTCATCTTCAAGCTGCGGSCAGAGAGAA 428  
 DB 570 CAGGCGAGTCTCTCAATTAAGACTCTCTATGCTGTGSCCAACTAGACACAGAGAGAA 629  
 OY 429 GATATGATGCTCATGACGACAGACACCCCACTCCAGTGTGTGTGTCACGTGTGAGAG 488  
 DB 630 GATATGATGCTCATGACGAGGAGTCAAGTCCACCCCAAGTGTGTGTGTCACGTGTGAGAG 689  
 OY 489 TGACCAAGAGATGAGAGAGA 509  
 DB 690 TGACATAGAGAGAGAGAGA 710

RESULT 13

ABL65187 standard; DNA; 3184 BP.

ABL65187;

15-MAY-2002 (first entry)

Lung cancer related gene sequence SEQ ID NO:3524.

Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 cytotoxic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 gene; db.

Homo sapiens.

WO200194629-A2.

13-DEC-2001.

30-MAY-2001; 2001WO-US10838.

05-JUN-2000; 2000US-209473P.

05-JUN-2000; 2000US-209531P.

18-SEP-2000; 2000US-233133P.

18-SEP-2000; 2000US-233617P.

20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 25-SEP-2000; 2000US-235637P.  
 PR 25-SEP-2000; 2000US-235638P.  
 PR 25-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237315P.  
 PR 02-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.

(AVAIL-) AVAILON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 Soppet DR, Weaver Z;  
 WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a  
 chemical agent to be tested for anti-neoplastic activity, and  
 determining a change in expression of a gene of a signature gene set  
 Claim 1; SEQ ID 3524; 44pp; English.

The present invention describes a method (M1) for screening for an  
 anti-neoplastic agent. The method involves exposing cells to a chemical  
 agent to be tested for anti-neoplastic activity, determining a change in  
 expression of at least one gene (I) of a signature gene set, where (I)  
 comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 to ABL70110), or is at least 95% identical to (S), where a change in  
 expression is indicative of anti-neoplastic activity. (I) has cytotoxic  
 activity and can be used in gene therapy. M1 can be used for screening  
 an anti-neoplastic agent, and can be used for producing a product which  
 is the data collected with respect to the anti-neoplastic agent as a  
 result of M1, and the data is sufficient to convey the chemical  
 structure and/or properties of the agent. M1 can be used in the  
 treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 carcinoma, papillary carcinoma and Wilm's tumour.

Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Query Match 35.9%; Score 214.2; DB 24; Length 3184;

Best Local Similarity 66.1%; Pred. No. 4.1e-56;



Matches 331; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

QY 21 CTACAAATTTAGTCTCTGATGCTGTGTGGCAACAGATGATGCTTTCAGCGAAAGTGA 80  
DB 210 CTGTGATGTTTCCACTCTGTGTCCTGTGTGGTGGATGTGAGGTCTTTACCAATCAGGA 269

QY 81 GACCAGGCGCAATTTGAATCCCTCTTCAGACATATGACAGGACACACCTTCCAGTA 140  
DB 270 GGTAAAGGAAATTTGGGGGACTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCT 329

QY 141 TTTTAAGAGCTTCAAAACGCTGTCGGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGC 200  
DB 330 ATTAAAGAGTTTCAGAGGTTGTCGTATAACTTCAGCAATCTTAATCTGCAGCCGAGC 389

QY 201 CAGGCTCGGCTGCACAAAGACGAGTTCTTGGGGAAGGAAATGAAGTGTATTTTCTCA 260  
DB 390 TAGGATAGAGCTTTCATGAACCACTTCAGAGGGGAAATTAAGCTCTTACTTTGCACA 449

QY 261 GACT-----TTACACATAGGAGTTTCACACCTGGCTCCGCCCAATCCCGACAA 308  
DB 450 GGTTCAGACTCCAGAGACAGATGGAGCAAACTGCACTTGGCTCCACCCAGCTGCCAA 509

QY 309 ACAGTTTCTCATCTCCCTCCGCGCTCTCTCCCTGGTGGCTGGAAACAAGTAGAAGATGC 368  
DB 510 ACAGTTTCTCATCTCCGCGCTCTCTCCCTGGTGGCTGGAAACAAGTAGAAGATGC 368

QY 369 CACCCCGCTCAATAATAGATCTTTATATGCACTCTCAAGCTGGGCGCCAGGAGAGAA 428  
DB 570 CAGCCGAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAGGACCGAGAGAA 629

QY 429 GTATGAACCTGCATGACGACGACACCCCACTCCAGTGTGGTGTCTCAGTGTGTGAG 488  
DB 630 GTATGAGCTCCATGCGAGGAGTGTGAGTCCACCCCAAGTGTGTGCTGTCGAGTGTGAG 689

QY 489 TGACCAAGAGATGAGGAGGA 509  
DB 690 TGACATAGAGGAGAGAGGA 710

RESULT 14  
ID AAF25338 standard; cDNA; 934 BP.  
XX AAF25338;  
AC AAF25338;  
DT 30-APR-2001 (first entry)  
XX Nucleotide sequence of a human detoxification protein.  
XX Human; detoxification protein; DTX; cancer; leukaemia; melanoma;  
KW adenocarcinoma; autoimmune disorder; inflammatory disorder;  
KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH 20..724  
CDS /\*tag= a  
/\*product= "detoxification protein"  
FT 20..100  
FT /\*tag= b  
FT sig\_peptide  
XX WO200104305-A2.  
XX 18-JAN-2001.  
XX 06-JUL-2000; 2000WO-US18509.  
XX 07-JUL-1999; 99US-0142678.  
XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H;  
PI WPI; 2001-147193/15.  
DR P-PSDB; AAB31788.  
XX New human detoxification protein and polynucleotide, useful for  
PT diagnosis, prevention and treatment of autoimmune/inflammatory  
PT disorders and cell proliferative disorders including cancer -  
XX Claim 5; Page 79; 79pp; English.

CC The present sequence encodes a human detoxification polypeptide (DTX).  
CC DTX and its (ant)agonists are useful for preventing or treating  
CC disorders associated with decreased or increased expression or activity  
CC of DTX. DTX polypeptides are useful for screening compounds that  
CC specifically binds to DTX and for identifying (ant)agonists.  
CC Diseases prevented, treated and diagnosed include cancers (e.g.  
CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,  
CC breast, kidney, liver, pancreas, prostate and uterus),  
CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,  
CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative  
CC colitis), bacterial, fungal, parasitic infections and cell  
CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,  
CC cirrhosis and hepatitis). Anti-DTX antibodies may be used as  
CC antagonists, as a targeting or delivery mechanism for bringing  
CC pharmaceutical agents into contact with cells or tissues expressing  
CC DTX and for diagnosis of DTX-related disorders.  
XX Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;

Query Match 35.1%; Score 209.6; DB 22; Length 934;  
Best Local Similarity 64.8%; Pred. No. 6.1e-55;  
Matches 333; Conservative 0; Mismatches 169; Indels 12; Gaps 1;

QY 8 TTAGGAGCTTTAGCTACAATTTTAGCTCCCTGATGCTGTGTGGCAACGATGATGCT 67  
DB 138 TCAGTACTTCAATGACTCCCACTCGTTGTTGGTGCATGTTTCACCACTCAGTGT 197

QY 68 TCAGCGAAAGTGAGACGAGCGGCAAAATTTGAATCCCTCTTCAGAACATATGACAAGACA 127  
DB 198 TTGAAGGAAGAGAGCAAGGAAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTG 257

QY 128 CCACCTTCCAGTATTTTAAGAGCTTCAAAACGCTGCCGATATAACTTCAGCAACCCCTTAT 187  
DB 258 TGAGGTTCCAGCTATTTAAGAGTTTCAGAGCTGTCGATATAACTTCAGCAATCCTAAAT 317

QY 188 CTGACGCCGATGCCAGCTGCGGCTGCCAAGACCGAGTTCCTGGGGAAGGAAATGAAGT 247  
DB 318 CTGACGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGGAAAAATTAAAGC 377

QY 248 TGTATTTTCTCAGACT-----TTACACATAGGAGTTTCACACCTCGCTCCGC 295  
DB 378 TCTACTTTCCAGGTTTCAGACTCCAGACAGATGAGACAACTGCACTTGGCTCCAC 437

QY 296 CCAATCCCGACAAACAGTTCTCATCTCCCTCCGCGCTCTCTCCCTGGTGGTGAAC 355  
DB 438 CCCAGCTGCCAAACAGTTTCTCATCTCGCCCTCTCTCTCCCTCCACCTGTTGGCTGCAGC 497

QY 356 AAGTAGAAGATGCCACCCCGCTCATATAAATAGATCTTTTATATGCTCATCTCCAAGTGG 415  
DB 498 CCATCAACGATGCCAGCGAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAG 557

QY 416 GGGCAGGAGAGATGATGAATGATGCTGACGACGACACCCCTCCAGTGTGTGCTGCC 475  
DB 558 GACCGAGAGAGATGATGAGCTCCATGCGAGGAGTGAAGTCCACCCCAAGTGTGCTGTC 617

QY 476 AGCTGTGTGAGAGTGACCAAGAGAAATGAGGAGGA 509  
DB 618 ACCTGTGCCAGCTGACATAGAGGAGAGAGGA 651

RESULT 15

ID	ABA91457
XX	ABA91457 standard; cDNA; 3159 BP.
XX	
AC	ABA91457;
XX	
DT	18-APR-2002 (first entry)
XX	
DE	Human Down syndrome critical region 1-like 1 protein cDNA.
XX	
KM	Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;
KW	Down syndrome; Alzheimer's disease; dementia; transgenic;
KM	neuroprotective; neotrophic; anticonvulsant; diagnosis;
XX	gene therapy; gene; aa.
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	185..952
FT	CDS
FT	/*tag= a
FT	/product= "DSCR1L1"
FT	/transl_except= (pos:215..217, aa:Xaa)
FT	/note= "Xaa = unknown"
XX	
PN	W0200204513-A2.
XX	
XX	17-JAN-2002.
PD	
XX	
PD	11-JUL-2001; 2001WO-US21982.
PF	
XX	11-JUL-2000; 2000US-0614474.
PR	
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
XX	Loring JF, Tingley DW, Edwards CM, Streeter DG;
XX	
DR	WPI; 2002-164633/21.
XX	P-PSDB; AAM50760.
PT	
PT	Novel Down syndrome critical region 1-like 1 protein and nucleic acid
PT	encoding the protein useful for diagnosis and treatment of Alzheimer's
PT	disease, Down syndrome and other forms of dementia -
XX	
PS	Claim 2; Page 46-47; 54pp; English.
XX	
CC	The present sequence is that of cDNA clone incyte ID No: 247500.5
CC	encoding novel Down syndrome critical region 1-like 1 protein
CC	(DSCR1L1 alpha, see AAM50760). Northern analysis indicated
CC	abundance in DSCR1L1 alpha in various libraries, with the highest
CC	abundance in tissues from the nervous system, including tissues
CC	associated with schizophrenia, Huntington's disease, epilepsy and
CC	amniotic lateral sclerosis. An absence of DSCR1L1 expression
CC	was observed in 7 of 8 libraries from subjects with Alzheimer's
CC	disease. A claimed method for detecting differential expression of
CC	a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down
CC	syndrome, Alzheimer's disease and other forms of dementia. A probe
CC	from such a nucleic acid is useful for identifying naturally
CC	occurring molecules encoding DSCR1L1 alpha, allelic variants or
CC	other molecules. The nucleic acids are useful for producing
CC	transgenic cell lines or organisms which model human disorders.
CC	They may also be used in gene therapy, and to screen for ligands
CC	which specifically bind the nucleic acid molecule, such as
CC	a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,
CC	peptide, transfection factor, repressor or regulatory molecule,
CC	for use as a therapeutic.

Db	366	TCACGACATTCGAATGACCTCCGCCAACTGTTGTTGGTCGAATGTTCAACGACATGT	425
Qy	68	TCAGCGAAGGAGACACGAGGCGCCAAATTTGATATCCCTCTCAGAACATAGCAAGACA	127
Db	426	TTGAAGGAAAGAGAGCAGGAAATAATTTAGGGAATGTTTCGACTTATGATGACTGTG	485
Qy	128	CCACCTTCACGATATTTTAAGAGCTTCAAGCGTGCCTCGGATPAACTTCAGCAACCCCTTAT	187
Db	486	TGACGTTCCAGCTATTATTAAGAGTTTTCAGACGTGTCCGTATTAATCTCAGCAATCTTAAT	545
Qy	188	CTGCGAGCCGATGCCAGGTGTGGGTGTGCACAAGATCCAGTTCCTTGGGGAAGAAATGAAT	247
Db	546	CTGCGAGCCCGAGCTTAGAGATAGAGCTTCATGAAACCCCAATTCAGAGGGAATAAATTAAGC	605
Qy	248	TGATATTTTGCTCAGACT-----TTACATATAGGAAGTTCACACTGAGCTCCGC	295
Db	606	TCTACTTTTGCACAGGTTTCAGACTCCAGACAGATGGAAGACAAATGCACTTGTGCTTCAC	655
Qy	296	CCAAATCCCGACAACAGTTCCTCATCTCCCTCCGACTCTCCCTCCCGTGGCTGGAAC	355
Db	666	CCCGAGCTCGCCAAACAGTTTCTCATCTCGCCCCCTCTCCCACTGTGGCTGGGAGC	725
Qy	356	AAGTAGAAGATGCCACCCCGCTCATTAATTAACATCTTTATATGTCATCTCCAAAGCTGG	415
Db	726	CCATCAACGATGCCACGCACTCTCAACTATGACTCTCTCTATGCTGTGGCCAAACTAG	785
Qy	416	GACCGAGAGAGAAATATGAATCTGCATSCAGGACAGACCCCACTCTCCAGATGTGTGCTCC	475
Db	786	GACCGAGAGAGAAATATGATGCTCATCAAGGACTGATGTCACCCCAAGTGTGTGTGTGC	845
Qy	476	ACGTGTGTGAGATGCCAACAGAAATAGAGGA	509
Db	846	ACGTGTGTGAGATGCCAATTAAGGAAGGA	879

Search completed: December 14, 2002, 15:19:01  
Job time : 102.497 secs

50 Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;  
 Query Match 35.1%; Score 209.6; DB 24; Length 3159;  
 Best Local Similarity 64.8%; Pred. NO. 1.1e-54;  
 Matches 333; Conservative 0; Mismatches 169; Indels 12; Gaps 1;  
 07 8 TTAGGACCTTGGCTACACATTTTAACTCCCTCATCTGTCTGTGTGGCAACGATGATCTCT 67  
 ||| |||||

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 1076.02 Seconds  
(without alignments)  
16146.861 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggatttttaggacttttag.....cacgataccaccttagctga 597

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues .

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_rod.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_hgo\_hum.\*

40: em\_hgo\_mus.\*

41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	597	6	AX365315 Sequence
2	597	100.0	597	10	AF237790 Mus muscu
3	590.6	98.9	2198	10	BC013551 Mus muscu
4	589	98.7	2141	10	AF263329 Mus muscu
5	515.8	86.4	597	6	AX365312 Sequence
6	515.8	86.4	597	10	AF237789 Mus muscu
7	511	85.6	621	10	AF282255 Mus muscu
8	511	85.6	2125	10	AF260717 Mus muscu
9	508.6	85.2	2224	10	AF263240 Mus muscu
10	504.4	84.5	599	6	AX365311 Sequence
11	504	84.4	626	10	AB075973 Rattus no
12	501	83.9	2216	10	CGU60263
13	470.6	78.8	2331	6	AX365321 Sequence
14	470.6	78.8	2346	9	HSU85267
15	459.6	77.0	2348	6	AX281651 Homo sapien
16	408.2	68.4	2212	6	AX365324 Sequence
17	408.2	68.4	2227	9	HSU85266
18	408.2	68.4	2289	9	BC002864
19	408.2	68.4	2407	9	AX092184
20	401.8	67.3	562	9	HSU53821
21	401.6	67.3	2284	9	HSU85265
22	399.6	66.9	2173	6	AX410694
23	399.6	66.9	2173	9	HSU28833
24	399	66.8	2679	9	AX055845
25	398	66.7	798	9	AF400429
26	398	66.7	2174	6	AR034241
27	214.2	35.9	3184	6	AX329596
28	214.2	35.9	3184	6	AX333015 Sequence
29	214.2	35.9	3184	6	AX365327 Sequence
30	214.2	35.9	3184	6	AX420436
31	214.2	35.9	3184	9	HUMZAK14
32	209.6	35.1	934	6	AX074351
33	209.6	35.1	3159	6	AX420425
34	209.6	35.1	3261	9	AK090990
35	206.4	34.6	3240	9	AY034085
36	206.4	34.6	3253	9	AY034086
37	201.4	33.7	594	6	AX365318
38	201.4	33.7	594	10	AF237791 Mus muscu
39	198.6	33.5	659	10	AF237887 Mus muscu
40	198.2	33.2	3227	10	AB061524
41	196.8	33.0	3295	10	AB061525
42	191.2	32.0	720	6	AX365333
43	191.2	32.0	720	9	AF176117
44	191.2	32.0	828	6	AX365330
45	191.2	32.0	828	6	AX420437

## ALIGNMENTS

RESULT 1  
AX365315  
LOCUS AX365315  
DEFINITION Sequence 5 from Patent WO0204491.  
ACCESSION AX365315  
VERSION AX365315.1  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Williams, S.R. and Rothermel, B.  
TITLE Methods and compositions relating to muscle selective calcineurin  
interacting protein (mcip)

597 bp DNA linear PAT 15-FEB-2002



```

Db 481 TGTGAGGTGACCAAGAGAAATGAGGAGGAGGAGAGATGAGAGAAATGAGAGACCC 540
QY 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597
Db 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597

RESULT 3
BC013551 2198 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, Down syndrome critical region homolog 1 (human),
DEFINITION clone MGC:19348 IMAGE:4236038, mRNA, complete cds.
ACCESSION BC013551
VERSION BC013551.1 GI:15488840
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 2198)
Direct Submission
Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapb-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 25 Row: m Column: 6.

FEATURES
source
1..2198
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19348 IMAGE:4236038"
/tissue_type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Kid14"
/lab_hosts="DH10B"
/notes="Vector: pCMV-SPORT6"
77..673
/codon_start=1
/product="Down syndrome critical region homolog 1 (human)"
/protein_id="AAH13551.1"
/db_xref="GI:15488841"
/db_xref="LocusID:54720"
/translation="MHFRDFSYNFSLLIACVANDVDFSEBSETRAKFESLFRTYDKDFTT
QYKFSKFRVIFNPSLGAADARLRLHKEFLGKEMKLYPACTLHIGSHLAPNPD
KQFLSPSPAPGVNKOVEDATPVINDYLLVALSKLGPCKEYELHNAITDTPSVVVHV
CESDQNEEEEMEMKPKKIIOTRPEYPIHLS"

BASE COUNT 533 a 539 c 569 g 557 t
ORIGIN

Query Match 98.9%; Score 590.6; DB 10; Length 2198;
Best Local Similarity 99.3%; Pred. No. 4.2e-160;
Matches 593; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGCATTTAGGACCTTAGCTACAATTTAGCTCCCTGATCTGTGTGGCAACGAT 60
Db 77 ATGCATTTAGGACCTTAGCTACAATTTAGCTCCCTGATCTGTGTGGCAACGAT 136

```

```

QY 61 GATGTCTTTCAGCGAAAGTGAAGACGAGGCCAAATTTGAATCCCTCTTTCAGAACATATGAC 120
Db 137 GATGTCTTTCAGCGAAAGTGAAGACGAGGCCAAATTTGAATCCCTCTTTCAGAACATATGAC 156
QY 121 AAGGACACCAACCTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGAATAAATCTTCAGCAAC 180
Db 197 AAGGACACCAACCTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGAATAAATCTTCAGCAAC 256
QY 181 CCCTTATCTGACGCGGATGCCAGGCTCGGCTGTCACAAGACCGAGTTCTCTGGGGAAGGAA 240
Db 257 CCCTTATCTGACGCGGATGCCAGGCTCGGCTGTCACAAGACCGAGTTCTCTGGGGAAGGAA 316
QY 241 ATGAAGTTGTATTTTCTGCTAGACTTTTACACATAGGAAGTTTACACCTGGTCCGCCCAAT 300
Db 317 ATGAAGTTGTATTTTCTGCTAGACTTTTACACATAGGAAGTTTACACCTGGTCCGCCCAAT 376
QY 301 CCGGACAAACAGTTCCTCATCTCCCTCCGGGCTCTCTCCCTGGCTGGGAAACAAGTA 360
Db 377 CCGGACAAACAGTTCCTCATCTCCCTCCGGGCTCTCTCCCTGGCTGGGAAACAAGTA 436
QY 361 GAAGATGCCACCCCGTCTATAATTACGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 420
Db 437 GAAGATGCCACCCCGTCTATAATTACGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 496
QY 421 GGAGAGAAGTATGAAGTGCATGTCAGCGACAGACCCACCTCCAGTGTGTGTCACAGTG 480
Db 497 GGAGAGAAGTATGAAGTGCATGTCAGCGACAGACCCACCTCCAGTGTGTGTCACAGTG 556
QY 481 TGTGAGATGACCAACAGAGAATGAGGAGGAGGAGAGATGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 557 TGTGAGATGACCAACAGAGAATGAGGAGGAGGAGAGATGAGAGAGATGAGAGAGATGAAGAGACCC 616
QY 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597
Db 617 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 673

RESULT 4
AF263239 2141 bp mRNA linear ROD 30-MAY-2000
LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
DEFINITION alternatively spliced.
ACCESSION AF263239
VERSION AF263239.1 GI:8102011
KEYWORDS Mus musculus.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 2141)
Fuentes J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1 transcripts
Unpublished
2 (bases 1 to 2141)
Fuentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
Direct Submission
Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
Avda. Castelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona
08907, Spain
Location/Qualifiers
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/notes="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"

FEATURES
source
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/notes="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"

```



TITLE \* A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling  
 JOURNAL J. Biol. Chem. (2000) In press  
 REFERENCE 2 (bases 1 to 597)  
 AUTHORS Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

FEATURES  
 source  
 1..597  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 1..597  
 /note="similar to Homo sapiens DSCR1; MC1P1; similar to EST sequence of GenBank Accession Number AA734360"  
 /codon\_start=1  
 /product="myocyte-enriched calcineurin interactin protein 1 splice variant 1"  
 /protein\_id="AAF63485.1"  
 /db\_xref="GI:7542526"  
 /translation="MEEVDLQDLPSATIACHLDRPVFDGLCRAKFESLFTYDKDTT FQYFKSRVRINFSNPLSADARLRLHKTFLGKMKLYFAOTLHGSHLAPPND KQFLISPPASPPVGMKQVEDATPVINDLLYAIKSGPGKYLHAATDTPSVVVHV CESQENEEBEEEMRMKPKKIOTRRPEYTPHLS"  
 BASE COUNT 160 a 170 c 151 g 116 t  
 ORIGIN

Query Match 86.4%; Score 515.8; DB 10; Length 597;  
 Best Local Similarity 94.4%; Pred. No. 1.8e-138;  
 Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 31 AGCTCCCTGATTGTTGTGGCAACGATGATGTTCTTCAGCGAAGTGAGACCGAGGCC 90  
 DB 31 AGCGCCACCATCGCTGCGACCGCGCGTGTTCGTGGAGCGCTGTGCGCGGCC 90  
 QY 91 AAATTGTAATCCCTCTTCAGAACATATGACAAGGACGACACCTTCAGTATTTAAAGC 150  
 DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAAGGACGACACCTTCAGTATTTAAAGC 150  
 QY 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTCGG 210  
 DB 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTCGG 210  
 QY 211 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270  
 DB 211 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270  
 QY 271 ATAGGAAGTTCCACCTGGCTCGGCCCAATCCCGACAAACAGTTCCTCATCTCCCTCG 330  
 DB 271 ATAGGAAGTTCCACCTGGCTCGGCCCAATCCCGACAAACAGTTCCTCATCTCCCTCG 330  
 QY 331 GCCTCTCTCCGTTGCTGGAAACAAGTAGAAGATGCCACCCCGCTGATAAATACGAT 390  
 DB 331 GCCTCTCTCCGTTGCTGGAAACAAGTAGAAGATGCCACCCCGCTGATAAATACGAT 390  
 QY 391 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAGTATGAACTGCATGACGAGCA 450  
 DB 391 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAGTATGAACTGCATGACGAGCA 450  
 QY 451 GACCCCACTCCCAAGTGTGTGCTCCAGCTGTGTGAGAGTGACCAAGAGATGAGGAGAA 510  
 DB 451 GACCCCACTCCCAAGTGTGTGCTCCAGCTGTGTGAGAGTGACCAAGAGATGAGGAGAA 510  
 QY 511 GAGGAGAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCAGACACGAGACCG 570  
 DB 511 GAGGAGAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCAGACACGAGACCG 570  
 QY 571 GAGTACACCGATCCACCTTACCTGA 597  
 DB 571 GAGTACACCGATCCACCTTACCTGA 597

RESULT 7  
 AF282255  
 LOCUS  
 DEFINITION  
 complete cds.  
 AF282255  
 VERSION  
 AF282255.1 GI:9652249  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 1 (bases 1 to 621)  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11080588  
 REFERENCE  
 2 (bases 1 to 621)  
 AUTHORS  
 TITLE  
 JOURNAL

FEATURES  
 source

1..621  
 /organism="Mus musculus"  
 /strains="BALB/c"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /sex="male"  
 /tissue\_type="brain"  
 1..621  
 /gene="Dscr1"  
 8..604  
 /gene="Dscr1"  
 /note="similar to the Homo sapiens Down syndrome candidate region 1 protein"

region 1 protein  
 /codon\_start=1  
 /product="Down syndrome candidate region 1 protein"  
 /protein\_id="AAF91461.1"  
 /db\_xref="GI:9652250"  
 /translation="MEEVDLQDLPSATIACHLDRPVFDGLCRAKFESLFTYDKDTT FQYFKSRVRINFSNPLSADARLRLHKTFLGKMKLYFAOTLHGSHLAPPND KQFLISPPASPPVGMKQVEDATPVINDLLYAIKSGPGKYLHAATDTPSVVVHV CESQENEEBEEEMRMKPKKIOTRRPEYTPHLS"  
 BASE COUNT 166 a 179 c 159 g 117 t  
 ORIGIN

Query Match 85.6%; Score 511; DB 10; Length 621;  
 Best Local Similarity 93.8%; Pred. No. 4.5e-137;  
 Matches 532; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
 QY 31 AGCTCCCTGATTGTTGTGGCAACGATGATGTTCTTCAGCGAAGTGAGACCGAGGCC 90  
 DB 31 AGCGCCACCATCGCTGCGACCGCGCGTGTTCGTGGAGCGCTGTGCGCGGCC 97  
 QY 91 AAATTGTAATCCCTCTTCAGAACATATGACAAGGACGACACCTTCAGTATTTAAAGC 150  
 DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAAGGACGACACCTTCAGTATTTAAAGC 157  
 QY 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTCGG 210  
 DB 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTCGG 217  
 QY 211 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270  
 DB 211 CTGCACAGACCGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 277  
 QY 271 ATAGGAAGTTCCACCTGGCTCGGCCCAATCCCGACAAACAGTTCCTCATCTCCCTCG 330



Db	278	ATAGAAATTACACCTGGCTCCGCCCAATCCCGAACAACAGTTCTCATCTCCCGCTCG	337
Qy	331	GCGCTCCCTCCCGTGTGGCTGGMAACAATGAGAATGCGACCCCGCTATATAATTAGAT	390
Db	338	GCGCTCTCTCCCGTGGCTGGMAACAATGAGAATGCGACCCCGCTATATAATTAGAT	397
Qy	391	CTTTTATATGCCATCTCCAACTGGGGCCAGAGAGAGATATGAACTCATATGCAAGCACA	450
Db	398	CTTTTATATGCCATCTCCAACTGGGGCCAGAGAGAGATATGAACTCATATGCAAGCACA	457
Qy	451	GACCCCACTCCCAAGTGTGGTGTCCACGTGTGTGAGTGAACCAAGAGATGAGAGAA	510
Db	458	GACACCACTCCCAAGTGTGGTGTCCACGTGTGTGAGTGAACCAAGAGATGAGAGAA	517
Qy	511	GAGGAAGAGATGAGAGAAATGAAGAACCCCAAGCCCAAAATCATCCACAACGGAGACCG	570
Db	518	GAGGAAGAGATGAGAGAAATGAAGAACCCCAAGCCCAAAATCATCCACAACGGAGACCG	577
Qy	571	GAGTACACACCGATCCACCTTAGCTGA	597
Db	578	GAGTACACACCGATCCACCTTAGCTGA	604

## RESULT 8

AF260717.1	2125 bp	mRNA	linear	ROD 22-MAR-2001
LOCUS	AF260717			
DEFINITION	Mus musculus Down syndrome candidate region 1 (Dscr1)	mRNA, complete		
ACCESSION	AF260717			
VERSION	AF260717.1			GI:7839596

ORGANISM

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2125)

Casas, C., Martinez, S., Pritchard, M.A., Fuentes, J.J., Nadal, M., Pineda, I., Soriano, E., Revilla, J. and

**TITLE**

expressed in the primitive ventricle of the hindbrain and is involved in neurogenesis  
Mech. Dev. 101 (1-2), 289-292 (2001)

## REFERENCE

AUTHORS	TITLE
Fuentes, J. J., Pritchard, M., Puchner, C. and Davidson, A.	Down syndrome candidate region 1 (Dscr1), one of three alternatively spliced exon 1

AUTHOR

**TITLE** Centre de Genetica Medica i Molecular, IRO  
**DIRECT SUBMISSION** Centre de Genetica Medica i Molecular, IRO  
**SUBMITTED (26-APR-2000)** Centre de Genetica Medica i Molecular, IRO  
**Avia, Castelldefels Km. 2,7, L'Hospitatet de Llobregat, Barcelona**

BOI

```

1 2125
/chromosome="16"
/db_xref="taxon:10090"
/Organism="Rae"

```

CD

FQYFKSFKRVRINFSNPLSADARLRHKTEFLGKEMKLYFAQTLHIGSSHLAPPNF

BASE COUNT	487	a	536	c	563	g	536	t	3	others
ORIGIN	KÖPLISBPASPRGKMQVEBATEVINDILLVATISKGRGKELHAATDTTPSVVAVH CESDÖNEEEEEEEMERKKRPXKIÖTRREPRYPIHLS"									

Query Match	85.64;	Score 511;	DB 10;	Length 2125;
Best Local Similarity	93.84;	Pred. No. 5,2e-117;		
Matches 532;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;

31 AGCTCCCGATTGCTTGTTGTTGGCAACGATGATGTCTTCACGGAAGTGACACCGAGGCC 90  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856

91 AAATTGATCCCTCTTCAGACATATGCACAAGACACCCTTCAGTATTTTAGACC 150

151 TTTAAACGCTGTCGGATAAAGCTCAGCAAGCCCTTATCTGCAAGCCGATGCGACAGGCTGCGG 210  
Db 105 AAATTTGATCCCTCTTCAGAACATATGACAAAGACACCACTCTCCAGTATTTTAAAGGC 164

Db 165 TTCAAACGTGTCGGATAACTTCAGCAACCCCTTATCTGACACCGATGCCAGGCTGCCG 224

Qy 211 CTCGACACGAACCGAGTCTCTGGGGAGGAAATGATGTTATTTTGTCTCAGACTTTCAC 270

Db 225 CTCGACACGAACCGAGTCTCTGGGGAGGAAATGATGTTATTTTGTCTCAGACTTTCAC 284

271 ATGGAAGTTCACACCTGCTCCGCCAATCCCGACAAAAGTTCTCATCTCCCTCCG 330

Db 285 ATGAGAGTTACACACTGGCTCCGCCCAATCCGACMAACAGTTCCTCATCTCCCTCCG 344

Oy 331 GCGCTCTCCCTCCGCTGCTGCGAAACAGTAGAAGATGCCACCCCGCTCATTAATTACGAT 390

Db 345 GCCTCTCCCTCCCGCTTGCGGAAACAGTAGAAGATGCCACCCCGTCATAAATTACGAT 404

Db	Sequence	Position
Oy	CTTTATATAGGCATCTCCAGCTGGGGCCAGAGAAATATGACATGATGACACAGACA	430
391	CTTTATATAGGCATCTCCAGCTGGGGCCAGAGAAATATGACATGATGACACAGACA	430
405	CTTTATATAGGCATCTCCAGCTGGGGCCAGAGAAATATGACATGATGACACAGACA	464

[illegible]

DB 465 GACACCACTCCAGATGTAAGTCCACGATGTGTAAGTACCACCGAAGTTGGCAGGTCCT  
QY 511 GAGAGAGATGGAGAATGAAGACCCAAAGCCAAAATCATCCAGACACGAGACCG 570

525 GAGGAGAGATGGAGGAATGAGACCCAAATTCATCCAGACCGGAGACCG 584

Uy 571 GAGTACACACCATCCACCTCAGCTGA 611  
 |||||  
 Db 585 GAGTACACACCATCCACCTCAGCTGA 611

RESULT 9  
A E263240

LOCUS	2224 bp	mRNA	linear	ROD	30-MAY-2008
DEFINITION	AP263240	Mus musculus calcineurin inhibitor	mRNA, complete cds,		
		alternatively spliced.			

ACCESSION	AF263240	GI:8102013
VERSION	AF263240.1	
KEYWORDS		

SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
1 (bases 1 to 2224)  
Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Eerivill,X.

**TITLE** Down syndrome candidate region 1 (Dschr1), one of three alternatively spliced exon 1 transcripts

**JOURNAL** Unpublished

REFERENCE  
AUTHORS  
TITLE  
2 (bases 1 to 2224)  
Fuentes, J. J., Pritchard, M. A., Pucharcos, C. and Estévil, X.  
Direct Submission

JOURNAL  
Submitted (02-MAY-2000)  
Centre de Genètica mèdica i Molecular,  
Avia. Castrelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona  
08907, Spain





Yokota, M.  
Direct Submission  
Submitted (11-DEC-2001) Koji Obata, Nagoya University, Department  
of Clinical Pathophysiology; 65 Tsubumi-cho, Showa-ku, Nagoya,  
Aichi 466-8550, Japan (E-mail: obata@nagoya-u.ac.jp,  
Tel: 81-52-744-2577, Fax: 81-52-744-2977)

FEATURES  
Location/Qualifiers

source

1. .626  
/organism="Rattus norvegicus"

gene  
/db\_xref="taxon:10116"

CDS  
1. .626  
/gene="mclp 1"

19. .618  
/gene="mclp 1"

/codon\_start=1  
/product="myocyte-enriched calcineurin-interacting protein  
1"

/protein\_id="BAC06443.1"

/db\_xref="GI:21998844"

/translation="MHPRDPMYNSPLIACVANGDVSESETRAKFESLPRTYDKDIT  
FOYFKSFKRVKINFSNPLSADARLRLHKEFELEKEMKLYFAOTLHIGSHLAPPD  
KOFILSPASPVGKQVEDATPVINDLVAISKLGEKTELHAATDTTPSVVHV  
CESDOEEEBEEMERMKRPKIIQTRREYPIHLIS"

BASE COUNT 182 a 163 c 154 g 127 t

ORIGIN  
Query Match 84.4%; Score 504; DB 10; Length 626;  
Best Local Similarity 91.1%; Pred. No. 4, 9e-135;  
Matches 547; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 1 ATGGATTTTAAAGGACTTGTAGCTACATTTTAACTTCCCTGCTGTGTGTGCAACGAT 60  
DB 19 ATGCATTTTAAAGGACTTGTAGCTACATTTTAACTTCCCTGCTGTGTGTGCAACGAT 78

QY 61 GATGCTTCAGCGAAAGTAGAGACCAAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 120  
DB 79 GATGCTTCAGCGAAAGTAGAGACCAAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 138

QY 121 AAGGACACCACTTCGAGATTTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 180  
DB 139 AAGGACATCACCTTCGAGATTTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 198

QY 181 CCCTTATCTGACCGATGCGAGCTCGGCTGCAAGACCGAGTTCTCGGGAAGAA 240  
DB 199 CCCTGCTGCGGAGGAGCGCGAGCTCGGCTGCAAGACCGAGTTCTCGGGAAGAA 258

QY 241 ATGAAGTTGATTTTGTCTGACCTTTACATAGGAAGTTCAACCTGCGTCCGCCAAT 300  
DB 259 ATGAAGTTGATTTTGTCTGACCTTTACATAGGAAGTTCAACCTGCGTCCGCCAAT 318

QY 301 CCGGACAAAGTTCCTCATCTCCCTCCGAGCTCTCTCCGTTGCTGGAACAAATA 360  
DB 319 CCGGACAAAGTTCCTCATCTCCCTCCGAGCTCTCTCCGTTGCTGGAACAAATA 378

QY 361 GAAGATCCACCCCGCTCATTAATTAAGATCTTTATATGCGATCTTCAAGCTGGGCCA 420  
DB 379 GAAGATCCACCCCGCTCATTAATTAAGATCTTTATATGCGATCTTCAAGCTGGGCCA 438

QY 421 GGAGAGAAGTTGAATGCTGACGAGCAAGACCCCACTCCAGTGTGTGTGCTCAAGT 480  
DB 439 GGAGAGAAGTTGAATGCTGACGAGCAAGACCCCACTCCAGTGTGTGTGCTCAAGT 498

QY 481 TGTGAGAGTACCAAGAGATGAGAGGAA--GAGGAAGATGAGAGAGATGAAGA 537  
DB 499 TGTGAGAGTACCAAGAGAGATGAGAGGAA--GAGGAAGATGAGAGAGATGAAGA 558

QY 538 CCCAAGCCCAAAATCATTCAGACAGGAGACCGGAGTACACACCGATTCACCTTAGCTGA 597  
DB 559 CCCAAGCCCAAAATCATTCAGACAGGAGACCGGAGTACACACCGATTCACCTTAGCTGA 618

RESULT 12  
CGU60263

LOCUS CGU60263 2216 bp mRNA linear ROD 03-SEP-1997  
DEFINITION Cricetus griseus Adapc78 (adapc78) mRNA, complete cds.

ACCESSION U60263

VERSION U60263.1

KEYWORDS GI:2351390

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

1. .2216  
/organism="Cricetus griseus"

/db\_xref="taxon:10029"

/cell\_type="HA-1"

1. .2216  
/gene="adapc78"

70. .663  
/gene="adapc78"

/note="similar to human DSCR1; an oxidant- and  
calcium-inducible mRNA"

/product="Adapc78"

/protein\_id="AB68517.1"

/db\_xref="GI:2351391"

/translation="MHPRDPMYNSPLIACVANGDVSESETRAKFESLPRTYDKDIT  
FOYFKSFKRVKINFSNPLSADARLRLHKEFELEKEMKLYFAOTLHIGSHLAPPD  
KOFILSPASPVGKQVEDATPVINDLVAISKLGEKTELHAATDTTPSVVHV  
CESDOEEEBEEMERMKRPKIIQTRREYPIHLIS"

BASE COUNT 544 a 515 c 564 g 593 t

ORIGIN  
Query Match 83.9%; Score 501; DB 10; Length 2216;  
Best Local Similarity 91.1%; Pred. No. 4, 1e-134;  
Matches 544; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 1 ATGGATTTTAAAGGACTTGTAGCTACATTTTAACTTCCCTGCTGTGTGTGCAACGAT 60  
DB 70 ATGCATTTTAAAGGACTTGTAGCTACATTTTAACTTCCCTGCTGTGTGTGCAACGAT 129

QY 61 GATGCTTCAGCGAAAGTAGAGACCAAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 120  
DB 130 GATGCTTCAGCGAAAGTAGAGACCAAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 189

QY 121 AAGGACACCACTTCGAGATTTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 180  
DB 190 AAGGACATCACCTTCGAGATTTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 249

QY 181 CCCTTATCTGACCGATGCGAGCTCGGCTGCAAGACCGAGTTCTCGGGAAGAA 240  
DB 250 CCCTTATCTGACCGATGCGAGCTCGGCTGCAAGACCGAGTTCTCGGGAAGAA 309

QY 241 ATGAAGTTGATTTTGTCTGACCTTTACATAGGAAGTTTCAACCTGCGTCCGCCAAT 300  
DB 310 ATGAAGTTGATTTTGTCTGACCTTTACATAGGAAGTTTCAACCTGCGTCCGCCAAT 369

QY 301 CCGGACAAAGTTCCTCATCTCCCTCCGAGCTCTCTCCGTTGCTGGAACAAATA 360  
DB 370 CCGGACAAAGTTCCTCATCTCTCCCTCCGAGCTCTCTCCGTTGCTGGAACAAATA 429





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 4201.35 Seconds  
(without alignments)

Title: US-09-782-953-11 X NOT MCI P  
 Perfect score: 2331  
 Sequence: 1 tttttttttccccggagg .....gggcacagatgctccttac 2331

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

[illegible]

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

```

GenEmbl.*
1: gb_ba.*
2: gb_ba.*
3: gb_htg.*
4: gb_in.*
5: gb_om.*
6: gb_ov.*
7: gb_pat.*
8: gb_ph.*
9: gb_pl.*
10: gb_pr.*
11: gb_ro.*
12: gb_scs.*
13: gb_sy.*
14: gb_un.*
15: gb_vl.*
16: em_ba.*
17: em_fun.*
18: em_hum.*
19: em_in.*
20: em_mu.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_scs.*
29: em_un.*
30: em_vl.*
31: em_htg_hum.*
32: em_htg_inv.*
33: em_htg_other.*
34: em_htg_mus.*
35: em_htg_pln.*
36: em_htg_rod.*
37: em_htg_man.*
38: em_htg_vrt.*
39: em_sy.*
40: em_htg_hum.*
41: em_htg_mus.*
42: em_htg_other.*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	2331	100.0	2331	6	AX365321	Sequence
2	2331	100.0	2346	9	HSU85267	U85267 Homo sapien
3	2318.4	98.5	2348	6	AX281651	AX281651 Sequence
3	2109	90.5	2212	6	AX365324	Sequence
5	2109	90.5	2227	9	HSU85266	U85266 Homo sapien
6	2107.4	90.4	2289	9	BC002864	BC002864 Homo sapi
7	2105.8	90.3	2407	9	AX092184	AX092184 Homo sapi
8	2104	90.3	2284	9	HSU85265	U85265 Homo sapien
9	2101	90.1	2173	6	AX411694	AX411694 Sequence
10	2101	90.1	2173	6	HSU28833	U28833 Homo sapien
11	2093.6	89.8	2679	9	AK055845	AK055845 Homo sapi
12	2059.2	88.3	2174	6	AR034241	AR034241 Sequence
C 13	1790.4	76.8	1792	9	AY007090	AY007090 Homo sapi
C 14	1767.4	75.8	83762	9	AP000326	AP000326 Homo sapi
C 15	1767.4	75.8	100000	9	AP000054	AP000054 Homo sapi
C 16	1767.4	75.8	100000	9	AP000054	AP000054 Homo sapi
C 17	1767.4	75.8	100000	9	AP000169	AP000169 Homo sapi
C 18	1767.4	75.8	100000	17	AP0000122	AP0000122 Homo sapi
C 19	1767.4	75.8	340000	9	AP001720	AP001720 Homo sapi
19	1011.8	43.4	2216	10	CGU60263	U60263 Cricetus
20	803.4	34.5	2198	10	BC013551	BC013551 Mus muscu
21	798.2	34.2	2141	10	AF263239	AF263239 Mus muscu
22	755.4	32.4	798	9	AF400423	AF400423 Homo sapi
23	720	30.9	2224	10	AF263240	AF263240 Mus muscu
24	714.6	30.7	2125	10	AB075971	AB075971 Mus muscu
25	552.2	23.7	562	9	HSU53821	U53821 Homo sapien
26	471.6	20.2	626	10	AB075973	AB075973 Rattus no
27	470.6	20.2	597	6	AX365315	AX365315 Sequence
28	470.6	20.2	597	10	AP237790	AP237790 Mus muscu
29	413.8	17.8	599	6	AX365311	AX365311 Sequence
30	412.2	17.7	621	10	AF282255	AF282255 Mus muscu
31	403.8	17.3	597	6	AX365312	AX365312 Sequence
32	403.8	17.3	597	10	AF237789	AF237789 Mus muscu
C 33	365.2	15.7	385	6	AX156214	AX156214 Sequence
C 34	261.8	11.2	303	11	G34765	G34765 human STS S
C 35	246.8	10.6	251	6	AX429944	AX429944 Sequence
36	221.6	9.5	90753	2	AC111585	AC111585 Rattus no
37	195.2	8.4	3227	10	AB061524	AB061524 Mus muscu
38	194.4	8.3	3184	6	AX329536	AX329536 Sequence
39	194.4	8.3	3184	6	AX333015	AX333015 Sequence
40	194.4	8.3	3184	6	AX365327	AX365327 Sequence
41	194.4	8.3	3184	6	AX420436	AX420436 Sequence
42	194.4	8.3	3184	9	HUMZAK74	D83407 ZAK1-4 mRNA
43	193.2	8.3	594	6	AX365318	AX365318 Sequence
44	193.2	8.3	594	10	AF237791	AF237791 Mus muscu
45	192.4	8.3	559	10	AF237787	AF237787 Mus muscu

## ALIGNMENTS

RESULT 1  
 AX365321  
 LOCUS  
 DEFINITION  
 Sequence 11 from Patent WO0204491.  
 ACCESSION  
 AX365321  
 VERSION  
 AX365321.1  
 GI:18697049  
 KEYWORDS  
 .  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 AUTHORS  
 Williams,S.R. and Rothermel,B.  
 TITLE  
 Methods and compositions relating to muscle selective calcineurin  
 interacting protein (mcip)

JOURNAL Patent: WO 0204491-A 11 17-JAN-2002;  
 Board of Regents, The University of Texas System (US) ; Williams,  
 Sanders R. (US) ; Rothenmel, Beverly (US)  
 Location/Qualifiers

FEATURES  
 Source 1. 2331  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 144. 737  
 /note="unamed protein product"

CDS  
 /codon\_start=1  
 /protein\_id="CAD23811.1"  
 /db\_xref="GI:18697050"  
 /translation="MHFANRYGSSSLIACVANSDFPSEETAKESLPRDYKDT  
 FOYKSPKRVINSPSADARLQHKIEFLKEMKLI PAQTLHIGSHLAPND  
 KQFLISPPSPVKQVEDATPVINDLAISKIDPGEKTELHMAITTPSVVHV  
 CESDKEEEMERMRPKKIIQTRPEVTPHLS"

BASE COUNT 630 a 470 c 547 g 684 t  
 ORIGIN

Query Match 100.0%; Score 2331; DB 6; Length 2331;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTTTCCTCCAGGAGTGGGGGCTGGCCCTTACTGCTTATTAAGACAGCTCAAGA 60  
 Db 1 TTTTTCCTCCAGGAGTGGGGGCTGGCCCTTACTGCTTATTAAGACAGCTCAAGA 60  
 Qy 61 AGGAACCTACAGCCTCTTGAAAGAAATCTACTAGGGGCTTACTGCTGCTGTG 120  
 Db 61 AGGAACCTACAGCCTCTTGAAAGAAATCTACTAGGGGCTTACTGCTGCTGTG 120  
 Qy 121 CGCTTCTACTGTAAAGAAAGAAATGATTTTGAACCTTAACTACAGTTTAACTCC 180  
 Db 121 CGCTTCTACTGTAAAGAAAGAAATGATTTTGAACCTTAACTACAGTTTAACTCC 180  
 Qy 181 TGATGCTCTGTGGCAACAGTATATCTTCAGCGAAGTGAACCAAGGCCAAATTG 240  
 Db 181 TGATGCTCTGTGGCAACAGTATATCTTCAGCGAAGTGAACCAAGGCCAAATTG 240  
 Qy 241 AGTCCCTCTTTAGAGCCTATGACAGACATCACCTTCAGTATTTTAAAGCTTCAAC 300  
 Db 241 AGTCCCTCTTTAGAGCCTATGACAGACATCACCTTCAGTATTTTAAAGCTTCAAC 300  
 Qy 301 GAGTCAGAAATTAACCTTACAGAACCCCTTCGCGCAGAGATGCGAGCTTCACTGATA 360  
 Db 301 GAGTCAGAAATTAACCTTACAGAACCCCTTCGCGCAGAGATGCGAGCTTCACTGATA 360  
 Qy 361 AGACTGAGTTTCTGGAAAGAAATGAAGTTATTTTGTCTCAGACCTTACACTAGGAA 420  
 Db 361 AGACTGAGTTTCTGGAAAGAAATGAAGTTATTTTGTCTCAGACCTTACACTAGGAA 420  
 Qy 421 GCTCAACCTTGGCTCCGCAAAATCCAGACAGGCTTCTGATCTCCCTCCGCTCTC 480  
 Db 421 GCTCAACCTTGGCTCCGCAAAATCCAGACAGGCTTCTGATCTCCCTCCGCTCTC 480  
 Qy 481 CGCCAGTGGATGGAACAAATGGAAGATGCGACCCCATTAATATATGATCTCTTAT 540  
 Db 481 CGCCAGTGGATGGAACAAATGGAAGATGCGACCCCATTAATATATGATCTCTTAT 540  
 Qy 541 ATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTCAGACGCGCATGACCA 600  
 Db 541 ATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTCAGACGCGCATGACCA 600  
 Qy 601 CTCCTCAGGCTGGTGTGCTCATGTATGTAGAGTATCAAGAGAGAGAGAGAGAA 660  
 Db 601 CTCCTCAGGCTGGTGTGCTCATGTATGTAGAGTATCAAGAGAGAGAGAGAGAA 660  
 Qy 661 TGGAAAGATGAGAGACCTTAAGCCAAATTTATCCAGACGAGAGGCGGAGTACAGC 720  
 Db 661 TGGAAAGATGAGAGACCTTAAGCCAAATTTATCCAGACGAGAGGCGGAGTACAGC 720  
 Qy 721 CGATCACCCTCAGCTGAATGTGGCAGCGCAGAGCGCTTCCAATCATCTCAGGGA 780

Db 721 CGATCACCCTCAGCTGAATGTGGCAGCGCAGAGCGCTTCCAATCATCTCAGGGA 780  
 Qy 781 GGAATCTTTTACTGTGGAGGTGGGCTGTGACACCTTCTCGAGAGTGGCAGCCAGATCG 840  
 Db 781 GGAATCTTTTACTGTGGAGGTGGGCTGTGACACCTTCTCGAGAGTGGCAGCCAGATCG 840  
 Qy 841 GGGTGGCAGAAATCCCAAGTTTCAATGTTGCTCAGAAAGAAATCAAGCCGTGTCCTTGT 900  
 Db 841 GGGTGGCAGAAATCCCAAGTTTCAATGTTGCTCAGAAAGAAATCAAGCCGTGTCCTTGT 900  
 Qy 901 CTATGCTGCAACCCAGTTACTGTTTCAATGGCAGCGGGAAATGACTTGGGCCAATGACTGA 960  
 Db 901 CTATGCTGCAACCCAGTTACTGTTTCAATGGCAGCGGGAAATGACTTGGGCCAATGACTGA 960  
 Qy 961 GTTTGGTGTATGCAACAAGACATTTGGAGCTGTCTTGAAGAAACAGATAATGATAGTG 1020  
 Db 961 GTTTGGTGTATGCAACAAGACATTTGGAGCTGTCTTGAAGAAACAGATAATGATAGTG 1020  
 Qy 1021 TTTTGTACTGTTCTTTTCTGTAGGTTCTGTCTGTGCAAGGGCAGTTGATCATGTAG 1080  
 Db 1021 TTTTGTACTGTTCTTTTCTGTAGGTTCTGTCTGTGCAAGGGCAGTTGATCATGTAG 1080  
 Qy 1081 CTCAGAGAGAGCTTCTGTTTCTAATGGGCTGCAAGGGCCACTCTTACTGTAGGAA 1140  
 Db 1081 CTCAGAGAGAGCTTCTGTTTCTAATGGGCTGCAAGGGCCACTCTTACTGTAGGAA 1140  
 Qy 1141 GAGGTACACAGAAAGCCGCTATGTCAGAGAGTTGTGAAACACAGACATGCAATGT 1200  
 Db 1141 GAGGTACACAGAAAGCCGCTATGTCAGAGAGTTGTGAAACACAGACATGCAATGT 1200  
 Qy 1201 GGAATTTGTAGCTTCTTCTTCTCCCTCAATGTTCTCATGTTGTGATATTAATG 1260  
 Db 1201 GGAATTTGTAGCTTCTTCTTCTCCCTCAATGTTCTCATGTTGTGATATTAATG 1260  
 Qy 1261 ATTTACAGACTAATCTTGTGTTGATTAAGTTACACCGTGTGTTTACATCTTT 1320  
 Db 1261 ATTTACAGACTAATCTTGTGTTGATTAAGTTACACCGTGTGTTTACATCTTT 1320  
 Qy 1321 GGAAGCCAGAAAGCGTTTGGAAACGATACCTTCCCAAGTTCTGGAGTCTGAC 1380  
 Db 1321 GGAAGCCAGAAAGCGTTTGGAAACGATACCTTCCCAAGTTCTGGAGTCTGAC 1380  
 Qy 1381 TCTTGGCAAGCAGCTTGTGCGGAACTCTCTGGAATGATCTCACTGACATCCCA 1440  
 Db 1381 TCTTGGCAAGCAGCTTGTGCGGAACTCTCTGGAATGATCTCACTGACATCCCA 1440  
 Qy 1441 ACCGTGCAACGTGTAATGTTGTTTGGCAAGAAAGTTGATCTGAAATCTCTGTA 1500  
 Db 1441 ACCGTGCAACGTGTAATGTTGTTTGGCAAGAAAGTTGATCTGAAATCTCTGTA 1500  
 Qy 1501 ATTTAGCTTATCAATTCAGAAATGAGAGTTTCACTGCCAATTTTATGAGTGAAGAA 1560  
 Db 1501 ATTTAGCTTATCAATTCAGAAATGAGAGTTTCACTGCCAATTTTATGAGTGAAGAA 1560  
 Qy 1561 TTTTGTAGTGTGTTTGGATGCGACCTCAATTTCTGTGTTCTTTTATGTTGTT 1620  
 Db 1561 TTTTGTAGTGTGTTTGGATGCGACCTCAATTTCTGTGTTCTTTTATGTTGTT 1620  
 Qy 1621 TCTATACATGAATCATAGCCAAATCTTTTGGAAACGTGTTGTTGATGATGTTT 1680  
 Db 1621 TCTATACATGAATCATAGCCAAATCTTTTGGAAACGTGTTGTTGATGATGTTT 1680  
 Qy 1681 TTTTACCCACAGAGACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATCTA 1740  
 Db 1681 TTTTACCCACAGAGACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATCTA 1740  
 Qy 1741 CTTGTGTACTTGGGTGAAGATATGCACTGGAGCTGAAGATGATTAACCTTACT 1800  
 Db 1741 CTTGTGTACTTGGGTGAAGATATGCACTGGAGCTGAAGATGATTAACCTTACT 1800  
 Qy 1801 GTGAATCATATGTTTGAAGAAAGCGTTCCCATGCTTAACAGACTTGAATTCAGAGCA 1860  
 Db 1801 GTGAATCATATGTTTGAAGAAAGCGTTCCCATGCTTAACAGACTTGAATTCAGAGCA 1860

QY 1861 TGCTCAAGTCGATAGTAGACTCTGTGGCGATATGAGAGGATGAGTGCCTTCCCAATTC 1920  
Db 1861 TGCTCAAGTCGATAGTAGACTCTGTGGCGATATGAGAGGATGAGTGCCTTCCCAATTC 1920  
QY 1921 TTCTCTGATGGAATGCTATATAGTGTAACTTGTAAATTTTCTAGTGTAAATGTT 1980  
Db 1921 TTCTCTGATGGAATGCTATATAGTGTAACTTGTAAATTTTCTAGTGTAAATGTT 1980  
QY 1981 ATGCTGTTAAATAGTATATATTTGGCCCTTACAATACCGTACCAATGTTTGTCAAT 2040  
Db 1981 ATGCTGTTAAATAGTATATATTTGGCCCTTACAATACCGTACCAATGTTTGTCAAT 2040  
QY 2041 TGAATATCTTAATGCAAGTAACTATGCTTTTGGAAATTTGGAAGATGTTTATTC 2100  
Db 2041 TGAATATCTTAATGCAAGTAACTATGCTTTTGGAAATTTGGAAGATGTTTATTC 2100  
QY 2101 TTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTTTCAATATCAAGAAATGATTC 2160  
Db 2101 TTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTTTCAATATCAAGAAATGATTC 2160  
QY 2161 AACGTTCTCAAAACCTGTTTACGCTTGTGTAAGAGGAGCCGCTTTGGGAGAGACAT 2220  
Db 2161 AACGTTCTCAAAACCTGTTTACGCTTGTGTAAGAGGAGCCGCTTTGGGAGAGACAT 2220  
QY 2221 TGCATCGCTGTCCTCAAGTGTCTTTCTGTTAAGTGTCTTTAACTGGAGGCTTAACCTCAAA 2280  
Db 2221 TGCATCGCTGTCCTCAAGTGTCTTTCTGTTAAGTGTCTTTAACTGGAGGCTTAACCTCAAA 2280  
QY 2281 ATACTTTTTTAACTGCACTTCTATAATAAATGGCCACAGATGCTCTTAC 2331  
Db 2281 ATACTTTTTTAACTGCACTTCTATAATAAATGGCCACAGATGCTCTTAC 2331

RESULT 2  
HSU85267  
LOCUS HSU85267 2346 bp mRNA linear PRI 19-APR-2000  
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,  
alternative exon 1, complete cds.  
ACCESSION U85267  
VERSION U85267.2 GI:7596915  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and Estivill,X.  
A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart  
Hum. Mol. Genet. 4 (10), 1935-1944 (1995)  
JOURNAL 96121593  
MEDLINE 8595418  
PUBMED 2 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M.A. and Estivill,X.  
Genomic organization, alternative splicing, and expression patterns of the DSCR1 (Down syndrome candidate region 1) gene  
Genomics 44 (3), 358-361 (1997)  
JOURNAL 97468152  
MEDLINE 9325060  
PUBMED 3 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M. and Estivill,X.  
Direct Submission  
TITLE Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain  
JOURNAL 97468152  
MEDLINE 9325060  
PUBMED 4 (bases 1 to 2346)  
Fuentes,J.J., Pritchard,M. and Estivill,X.  
Direct Submission  
TITLE Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain

REMARK Sequence update by submitter  
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612867.  
FEATURES  
Location/Qualifiers  
source 1..2346  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1-q22.2"  
gene 1..2346  
/gene="DSCR1"  
CDS 144..737  
/gene="DSCR1"  
/note="down syndrome candidate region 1; one of four alternatively spliced exon 1"  
/codon\_start=1  
/protein\_id="BAB84372.1"  
/translation="MHPNPNVSPSSLIACVANSDFSESETRAKFBSLFRYDKDIT  
FOYFKSFXRVIRNFNSPFSADARLQLHKTFGLKREMKLYFAQLHIGSSHLAPNP  
KQFLISPPASPPVGWKQVEDATPVNIYDLLYAISKLGPGEKLYELHAATDTTPSVVHV  
CESDOKEEBEBEEMRRPKKIIOTRPEYTPHLS"  
BASE COUNT 645 a 470 c 547 g 684 t  
ORIGIN  
Query Match 100.0%; Score 2331; DB 9; Length 2346;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTTTCCTTCCCGAGGAGTGGGGCTGCTTACTTCTTATAGCAGCAGCTCAAGA 60  
Db 1 TTTTTCCTTCCCGAGGAGTGGGGCTGCTTACTTCTTATAGCAGCAGCTCAAGA 60  
QY 61 AGGAACCTACAGCTCTTGGAAAGGAATCTCACTAGGGGCTTGACTGGGCTGTAG 120  
Db 61 AGGAACCTACAGCTCTTGGAAAGGAATCTCACTAGGGGCTTGACTGGGCTGTAG 120  
QY 121 CGCTTTCACTGTAAAGAGCAAGATGCAATTTTAGAACTTTAACTACAGTTTACCTCCC 180  
Db 121 CGCTTTCACTGTAAAGAGCAAGATGCAATTTTAGAACTTTAACTACAGTTTACCTCCC 180  
QY 181 TGATTGCTGTGTGCAACAGTGTATATCTTTCAGCGAAGTGAAACAGGCGCAATTTG 240  
Db 181 TGATTGCTGTGTGCAACAGTGTATATCTTTCAGCGAAGTGAAACAGGCGCAATTTG 240  
QY 241 AGTCCTCTTTAGGACGCTATCAGCAGGACATCAGCTTTAGTATTTTAAAGAGCTTCAAC 300  
Db 241 AGTCCTCTTTAGGACGCTATCAGCAGGACATCAGCTTTAGTATTTTAAAGAGCTTCAAC 300  
QY 301 GAGTCAGATAAATCTCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGCTGCATA 360  
Db 301 GAGTCAGATAAATCTCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGCTGCATA 360  
QY 361 AGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 420  
Db 361 AGACTGAGTTTCTGGAAAGGAATGAAGTTATATATTTTGTCTCAGACCTTACACATAGGAA 420  
QY 421 GCTCAGACCTGCTCCGCAAAATCCAGCAGGAGTTTCTGATCTCCCTCCCGCTCTC 480  
Db 421 GCTCAGACCTGCTCCGCAAAATCCAGCAGGAGTTTCTGATCTCCCTCCCGCTCTC 480  
QY 481 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCAATAAATATGATCTCTTAT 540  
Db 481 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCAATAAATATGATCTCTTAT 540  
QY 541 ATGCCATCTCCAGCTGGGCGCAGGGAAGATGATGATTTGACGAGCACTGACACCA 600  
Db 541 ATGCCATCTCCAGCTGGGCGCAGGGAAGATGATGATTTGACGAGCACTGACACCA 600  
QY 601 CTCCAGCGTGGTGGTCCATGATGTGTGAGAGTGTATCAAGAGAGGAGGAGAGAGAAA 660  
Db 601 CTCCAGCGTGGTGGTCCATGATGTGTGAGAGTGTATCAAGAGAGGAGGAGAGAGAAA 660  
QY 661 TGGAAAGATGAGGAGACCTTAAGCCAAATTTATCCAGACAGGAGCGCGGAGTACACGC 720



Db 661 TGGAAAGATGAGGAGCTTAAGCCAAATATATCAAGCCAGGAGGCGGAGATCAAGC 720  
 Qy 721 CGATCCACCTCAGCTGAA CTGGCAGCGA CGAGAGCGATTCGAATCAATCTCAAGGA 780  
 Db 721 CGATCCACCTCAGCTGAA CTGGCAGCGA CGAGAGCGATTCGAATCAATCTCAAGGA 780  
 Qy 781 GGAATCTTTTACTGTGAGAGTGGCTGTGTCAGCACTTCTCGAGAGTGGCAGCGGATCG 840  
 Db 781 GGAATCTTTTACTGTGAGAGTGGCTGTGTCAGCACTTCTCGAGAGTGGCAGCGGATCG 840  
 Qy 841 GGGTGGCAGAAATCCAGTTCAATGTTGCTCAGAGAGAAATCAAGGCCGTGTCCTTGT 900  
 Db 841 GGGTGGCAGAAATCCAGTTCAATGTTGCTCAGAGAGAAATCAAGGCCGTGTCCTTGT 900  
 Qy 901 CTAAATCTGACACACAGTTACTGTTCAATGCGACCCGGGAAATGACTTTGGCCAAATCACTGA 960  
 Db 901 CTAAATCTGACACACAGTTACTGTTCAATGCGACCCGGGAAATGACTTTGGCCAAATCACTGA 960  
 Qy 961 GTTTGTGTGATGCGCAAGAGACATTTGGAGACTGTGTAAGAAAACAGATATGATG 1020  
 Db 961 GTTTGTGTGATGCGCAAGAGACATTTGGAGACTGTGTAAGAAAACAGATATGATG 1020  
 Qy 1021 TTTTGTACTTGTCTTTCTGTGAGTGTCTGTCTGTGCAAGGGCAGTTGATCAAGTAG 1080  
 Db 1021 TTTTGTACTTGTCTTTCTGTGAGTGTCTGTCTGTGCAAGGGCAGTTGATCAAGTAG 1080  
 Qy 1081 CTCAGAGAGAGCTTCTGTTTCTTAAGTGGCTTCAGGGGCCCTCTCTACTGTATGAGAA 1140  
 Db 1081 CTCAGAGAGAGCTTCTGTTTCTTAAGTGGCTTCAGGGGCCCTCTCTACTGTATGAGAA 1140  
 Qy 1141 GAGGTCCACAGAGAGCGGCTGAGTGGAGAGTGTGAAAACAGACAAATGCAATGT 1200  
 Db 1141 GAGGTCCACAGAGAGCGGCTGAGTGGAGAGTGTGAAAACAGACAAATGCAATGT 1200  
 Qy 1201 GGAATTTGATGCGTTTCTTCTCCCTCAATGTTCTCATGTTGTGATATATATATCTG 1260  
 Db 1201 GGAATTTGATGCGTTTCTTCTCCCTCAATGTTCTCATGTTGTGATATATATATCTG 1260  
 Qy 1261 ATTTACAGACTAATCTTGTGTAATATATATATATATATATATATATATATATATCTG 1320  
 Db 1261 ATTTACAGACTAATCTTGTGTAATATATATATATATATATATATATATATATATCTG 1320  
 Qy 1321 GGGAGCCAGAGAAAGCGTTTGGAAAACGATATCACTTTCCAGATTCGAGATTCGAG 1380  
 Db 1321 GGGAGCCAGAGAAAGCGTTTGGAAAACGATATCACTTTCCAGATTCGAGATTCGAG 1380  
 Qy 1381 TCTTTGCAAGCAGCTGTTGCGGAACTCTTCTGGAATGCAATTCCTAGCATCCCA 1440  
 Db 1381 TCTTTGCAAGCAGCTGTTGCGGAACTCTTCTGGAATGCAATTCCTAGCATCCCA 1440  
 Qy 1441 ACCGTGCAACCTGTAATCTGTTGCAAAAAGAGTGAATCTGTAATCTCTGTAGA 1500  
 Db 1441 ACCGTGCAACCTGTAATCTGTTGCAAAAAGAGTGAATCTGTAATCTCTGTAGA 1500  
 Qy 1501 ATTAGCTTATACAAATTCAGAAATAGCAGTTTCACTGCCAACTTTAGTGGGTGAGAAA 1560  
 Db 1501 ATTAGCTTATACAAATTCAGAAATAGCAGTTTCACTGCCAACTTTAGTGGGTGAGAAA 1560  
 Qy 1561 TTTTGTAGTGTGTTGGATGCGGACCTGATTTCTGTTCTTTTAAATGAGTGGT 1620  
 Db 1561 TTTTGTAGTGTGTTGGATGCGGACCTGATTTCTGTTCTTTTAAATGAGTGGT 1620  
 Qy 1621 TCTATACATGAATATACCAAAAATTTTGGAAAATGTTGGTGAATAGTGGTTC 1680  
 Db 1621 TCTATACATGAATATACCAAAAATTTTGGAAAATGTTGGTGAATAGTGGTTC 1680  
 Qy 1681 TTTTACCCCAAGAGACATCAAGATACACTGTAAATTAAGCTGATATATATATCA 1740  
 Db 1681 TTTTACCCCAAGAGACATCAAGATACACTGTAAATTAAGCTGATATATATATCA 1740  
 Qy 1741 CCTGTGTACACTTGGGTGAAAAGTATGCAATGGGAGATTAAGTATTAACCTACT 1800

Db 1741 CCTGTGTACACTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGATTAACCTACT 1800  
 Qy 1801 GTGAATCATATGTTGTAGGAAAAGCTGTTCCATGCTTAACAGAGACTTGAATTCAAAGCA 1860  
 Db 1801 GTGAATCATATGTTGTAGGAAAAGCTGTTCCATGCTTAACAGAGACTTGAATTCAAAGCA 1860  
 Qy 1861 TGTCAAGTGTATAGTATGATCTGTGGCAATATGAGAGGATGCAAGTCCCTTCCCATTTCA 1920  
 Db 1861 TGTCAAGTGTATAGTATGATCTGTGGCAATATGAGAGGATGCAAGTCCCTTCCCATTTCA 1920  
 Qy 1921 TTCCGTATGGAATTTGTAATCTAGGTTAACTTGTATATTTTCTAGTGAATGTGT 1980  
 Db 1921 TTCCGTATGGAATTTGTAATCTAGGTTAACTTGTATATTTTCTAGTGAATGTGT 1980  
 Qy 1981 ATGTCTGTAAATAGTATATATATTTTGGCTTACAAATCCGTAAACAATGTTGTCAATT 2040  
 Db 1981 ATGTCTGTAAATAGTATATATATTTTGGCTTACAAATCCGTAAACAATGTTGTCAATT 2040  
 Qy 2041 TGAATATCTTAATGCCAAGTAAATGCAATGCAATGCTTTGGAAAATTTGGAAGATGTTTATTC 2100  
 Db 2041 TGAATATCTTAATGCCAAGTAAATGCAATGCAATGCTTTGGAAAATTTGGAAGATGTTTATTC 2100  
 Qy 2101 TTTGAGAGCAAAATATATGTTGCAATTAATGCTTGAATTTGTTCAATCAAGAAATTTGATTG 2160  
 Db 2101 TTTGAGAGCAAAATATATGTTGCAATTAATGCTTGAATTTGTTCAATCAAGAAATTTGATTG 2160  
 Qy 2161 AACGTTCTCAAAACCTGTTTACGGTACTGTGTAAGAGGAGCGGTTTGGAGAGACAT 2220  
 Db 2161 AACGTTCTCAAAACCTGTTTACGGTACTGTGTAAGAGGAGCGGTTTGGAGAGACAT 2220  
 Qy 2221 TGCATCGCTGTCCAAAGTCTTCTGTTTAAAGTCTTTTAACTGGAGAGGCTTAACCTCAAA 2280  
 Db 2221 TGCATCGCTGTCCAAAGTCTTCTGTTTAAAGTCTTTTAACTGGAGAGGCTTAACCTCAAA 2280  
 Qy 2281 ATACTTTTATTAATGCAATCTATATATATATGAGCAGATATGCTCTTAC 2331  
 Db 2281 ATACTTTTATTAATGCAATCTATATATATATGAGCAGATATGCTCTTAC 2331

RESULT 3  
 AX281651  
 LOCUS 2348 bp DNA linear PAT 02-NOV-2001  
 DEFINITION Sequence 60 from Patent WO0177389.  
 ACCESSION AX281651  
 VERSION AX281651.1 GI:16608902  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Shiffman, D., Somogyi, R., Lawn, R., Sellhammer, J. J., Porter, G. J.,  
 Mikita, T. and Tai, J.  
 TITLE Genes expressed in foam cell differentiation  
 JOURNAL Patent: WO 0177389-A 60 18-OCT-2001;  
 Incyte Genomics, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..2348  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 042176-5"  
 BASE COUNT 638 a 473 c 550 g 687 t  
 ORIGIN  
 Query Match 99.5%; Score 2318.4; DB 6; Length 2348;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2330; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 Qy 1 TTTTATTTTCCCGAGGAGATGGGGCTGGCCCTTACTGCTTATTAAGCAGCTCAAGA 60  
 Db 8 TTTTATTTTCCCGAGGAGATGGGGCTGGCCCTTACTGCTTATTAAGCAGCTCAAGA 67  
 Qy 61 AGGAACCTACAGCCTCTTGAAAAGAAATCTCACTAGGGGCTTGAATGCGTGTCTGAG 120



Db 68 AGGAACCTAAGCCTCTTGGAAGGAATCTCACTAGGGCTTGACTGGTGGGTCTGTAG 127  
QY 121 CGCTTTCACTGTAAAGAACAAAGATGCAATTTTAGAAACTTTTAACATACAGTTTGTAGTCC 180  
Db 128 CGCTTTCACTGTAAAGAACAAAGATGCAATTTTAGAAACTTTTAACATACAGTTTGTAGTCC 187  
QY 181 TGATGGCTGTGTGCGCAACAGTATATCTTCAGCGAAAGTGAACACAGGCGCAAAATTTG 240  
Db 188 TGATGGCTGTGTGCGCAACAGTATATCTTCAGCGAAAGTGAACACAGGCGCAAAATTTG 247  
QY 241 AGTCCCTCTTTAGGACGTATCAAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAC 300  
Db 248 AGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAC 307  
QY 301 GAGTCAGAAATACTTCAGCAACCCCTTCTCCGACAGAGATGCCAGGCTCCAGCTGCATA 360  
Db 308 GAGTCAGAAATACTTCAGCAACCCCTTCTCCGACAGAGATGCCAGGCTCCAGCTGCATA 367  
QY 361 AGACTGAGTTTCTGGGAAGGAATGAATTTATTTTGTCTCAGACCTTACATAGGAA 420  
Db 368 AGACTGAGTTTCTGGGAAGGAATGAATTTATTTTGTCTCAGACCTTACATAGGAA 427  
QY 421 GCTCACACT-GGCTCCGCCAAATCCAGACAGATTTCTGATCTCCCTCCCGCCTCT 479  
Db 428 GCTCACACTGGGCTCCGCCAAATCCAGACAGATTTCTGATCTCCCTCCCGCCTCT 487  
QY 480 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCATAAACTATGATCTCTTA 539  
Db 488 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCATAAACTATGATCTCTTA 547  
QY 540 TATGCCATCTCCAGCTGGGCGAGGGAAGATGAATTTGTCAGCGAGCGATGACACC 599  
Db 548 TATGCCATCTCCAGCTGGGCGAGGGAAGATGAATTTGTCAGCGAGCGATGACACC 607  
QY 600 ACTCCAGCTGGTGGTCCATGTATGTGAGAGTGAATCAAGAGAGGAGGAGAGAGAA 659  
Db 608 ACTCCAGCTGGTGGTCCATGTATGTGAGAGTGAATCAAGAGAGGAGGAGAGAGAA 667  
QY 660 ATGGAAGAGATGAGGACCTAAGCCAAAATTTATCCAGACCGAGGCGCGAGTACAG 719  
Db 668 ATGGAAGAGATGAGGACCTAAGCCAAAATTTATCCAGACCGAGGCGCGAGTACAG 727  
QY 720 CCGATCCACCTCAGCTCAACTGGCACCGCAGGAGCGCATTCGAATCATCTCAGGG 779  
Db 728 CCGATCCACCTCAGCTCAACTGGCACCGCAGGAGCGCATTCGAATCATCTCAGGG 787  
QY 780 AGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGGAGGTGGCAGCCGAGATC 839  
Db 788 AGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGGAGGTGGCAGCCGAGATC 847  
QY 840 GGGTGGCAGAAATCCAGTTTCATGTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGT 899  
Db 848 GGGTGGCAGAAATCCAGTTTCATGTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGT 907  
QY 900 TCTAATGCTGCACACCACTGTTCTATGGCACCCGGAAATGACITGGCCCAATCACTG 959  
Db 908 TCTAATGCTGCACACCACTGTTCTATGGCACCCGGAAATGACITGGCCCAATCACTG 967  
QY 960 AGTTTGTGGTATCGACAAAGGACATTTGGGACTGTCTTGAGAAAACAGATATGATGT 1019  
Db 968 AGTTTGTGGTATCGACAAAGGACATTTGGGACTGTCTTGAGAAAACAGATATGATGT 1027  
QY 1020 GTTTTGTACTTGTCTTTCTGTGTAGTGTCTGTCTGCGAAGGAGGAGTTGATCAGTGA 1079  
Db 1028 GTTTTGTACTTGTCTTTCTGTGTAGTGTCTGTCTGCGAAGGAGGAGTTGATCAGTGA 1087  
QY 1080 GCTCAGAGAGAGCTTCTGTCTTAAAGTGGCTGCGAGGGCCACTCTCTAATGTTAGGA 1139  
Db 1088 GCTCAGAGAGAGCTTCTGTCTTAAAGTGGCTGCGAGGGCCACTCTCTAATGTTAGGA 1147  
QY 1140 AGAGGTACCAAGGAGCCCTAGTGGCAGAGAGGTTGTGAAAACAGCAGCAATGCAATG 1199

Db 1148 AGAGTACCAACAGGAGCCCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATG 1207  
QY 1200 TGGAAATTTGAGCGTTTCTTTTCCCTCATGTTCTCATGTTTGTGCAATATATTAAT 1259  
Db 1208 TGGAAATTTGAGCGTTTCTTTTCCCTCATGTTCTCATGTTTGTGCAATATATTAAT 1267  
QY 1260 GATTTTCAAGACATAACCTTGTTCGTATATAAAGTTTACACCGTTGTGTTTACATCTTT 1319  
Db 1268 GATTTTCAAGACATAACCTTGTTCGTATATAAAGTTTACACCGTTGTGTTTACATCTTT 1327  
QY 1320 TGGGAAGCCAGGAAACGGTTTGGAAAACGTTATCACCTTTCCAGATTTCTCGAATCTCGA 1379  
Db 1328 TGGGAAGCCAGGAAACGGTTTGGAAAACGTTATCACCTTTCCAGATTTCTCGAATCTCGA 1387  
QY 1380 CTCTTTGCAACAGCACCTTCTGCGGAACTCTTCTGGAATGCAATTCAGTCAGCATCCCC 1439  
Db 1388 CTCTTTGCAACAGCACCTTCTGCGGAACTCTTCTGGAATGCAATTCAGTCAGCATCCCC 1447  
QY 1440 AACCGTGCACACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGMAATTTCTCTGTAG 1499  
Db 1448 AACCGTGCACACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGMAATTTCTCTGTAG 1507  
QY 1500 AATTTAGCTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTGTGCGGTGAGAA 1559  
Db 1508 AATTTAGCTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTGTGCGGTGAGAA 1567  
QY 1560 AATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGT 1619  
Db 1568 AATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGT 1627  
QY 1620 TTCTATACATGAATCATAGCCAAACCTTTTGGAAAACCTTTTGGAAAACCTTTTGTGAGTAGTGTGTT 1679  
Db 1628 TTCTATACATGAATCATAGCCAAACCTTTTGGAAAACCTTTTGTGAGTAGTGTGTT 1687  
QY 1680 CTTTATACCCACGAAAGACATCAAGATACACTTGTAAAATAAGCTATAGCATATATTCAT 1739  
Db 1688 CTTTATACCCACGAAAGACATCAAGATACACTTGTAAAATAAGCTATAGCATATATTCAT 1747  
QY 1740 ACTGTTGTACACTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTATTAACCTACC 1799  
Db 1748 ACTGTTGTACACTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTATTAACCTACC 1807  
QY 1800 TGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGTTCTAACAGGACTTGAATTTCAAAGC 1859  
Db 1808 TGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGTTCTAACAGGACTTGAATTTCAAAGC 1867  
QY 1860 ATGTCAGTGGATAGTATCTGTGGCGATATGAGAGGATGCGAGTGCCTTTCCCAATTC 1919  
Db 1868 ATGTCAGTGGATAGTATCTGTGGCGATATGAGAGGATGCGAGTGCCTTTCCCAATTC 1927  
QY 1920 ATTCTGATGAAATTTGTTATATACTAGGTTTAACTTTGTAATTTTCTTAGTTGTAATGTG 1979  
Db 1928 ATTCTGATGAAATTTGTTATATACTAGGTTTAACTTTGTAATTTTCTTAGTTGTAATGTG 1987  
QY 1980 TATGTCGTGTAATAGGTAATTTATTTTGGCTTACAAATACCGTAACAAATGTTTGTCAAT 2039  
Db 1988 TATGTCGTGTAATAGGTAATTTATTTTGGCTTACAAATACCGTAACAAATGTTTGTCAAT 2047  
QY 2040 TTTGAAATACCTTAATGCCAAGTAACATGCGATGCTTTGGAAATTTGGAGATGTTTATT 2099  
Db 2048 TTTGAAATACCTTAATGCCAAGTAACATGCGATGCTTTGGAAATTTGGAGATGTTTATT 2107  
QY 2100 CTTTGGAGAAGCAAAATATGTTTGGCAATTAATGCTTTTGAATGTTTATATCAAGAAATTCAT 2159  
Db 2108 CTTTGGAGAAGCAAAATATGTTTGGCAATTAATGCTTTTGAATGTTTATATCAAGAAATTCAT 2167  
QY 2160 GAACGTTTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCA 2219  
Db 2168 GAACGTTTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCA 2227  
QY 2220 TTGCATCGCTGCTCAAGTGTCTTCTGTTAGTGTCTTTTAACTGGAGAGGCTAACCTCAA 2279  
Db 2228 TTGCATCGCTGCTCAAGTGTCTTCTGTTAGTGTCTTTTAACTGGAGAGGCTAACCTCAA 2287

QY	2280	AACTACTTTTAACTGATCTCTATTAATTAATGGGACAGATCTCTTAC	2331
Db	2288	AAATATTTTAACTGATCTCTATTAATTAATGGGACAGATCTCTTAC	2339
RESULT 4			
AX365324	AX365324	2212 bp	DNA
LOCUS	Sequence 14 from Patent WO2004491.		linear
DEFINITION	AX365324		PAT 15-FEB-2002
ACCESSION	AX365324		
VERSION	AX365324.1		
KEYWORDS	GI:18697051		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1		
AUTHORS	Williams S.R. and Rothermel B.		
TITLE	Methods and compositions relating to muscle selective calcineurin		
	intersecting protein (mip1)		
JOURNAL	Patent: NO 0204491-A 14 17-JAN-2002;		
	Board of Regents, The University of Texas System (US) ; Williams,		
	Sanders R. (US) ; Rothermel, Beverly (US)		
FEATURES	Location/Qualifiers		
source	1..2212		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
CDS	25..618		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD23812.1"		
	/db_xref="GI:18697052"		
	/translation="MEEVLDLDDLEPATIACHLDBRVFVDGLCRAKFESLFRDYDIT		
	ROYFSPKRPVRIINSPSADARLQKLEFLGEMKLPAQTLIGSSHLAPNDP		
	KQFLISPPASPVGVKQVEDATPVNVLKLTALVLSIGKEGXEHLAATDTTPSVVAV		
	CESDQKEEEMEMRRPKFIOTRKPETPHLS"		
BASE COUNT	588 a 455 c 529 g 640 t		
ORIGIN			
Query Match	90.5%	Score 2109;	DB 6; Length 2212;
Best Local Similarity	97.5%	Pred. No. 0;	
Matches 2142; Conservative	0;	Mismatches 55;	Indels 0; Gaps 0;
QY	135	GAAGCAAGATGCATTTAGAACTTTAACTTACATACAGTTTAACTCTCCTGATTCCTGCTGTG	194
Db	16	GACTGCGAAGTGGAGAGAGGTGACCTGCGACAGCTCGCCAGCCCACTCGCTGTGAC	75
QY	195	GCAAAACATGATATCTTACCGCGAAAGTGAACACAGGCGCAAATTAAGTCCCTTTAG	254
Db	76	CTGACACCCGCGCGTTCGTGACCGCGCTGTGCGGCGCAAATTAAGTCCCTTTAG	135
QY	255	ACGTATGACAGAGCATCACCTTTCAGTATTTAAGAGCTTCAAAAGAGTCAGATTAAC	314
Db	136	ACGTATGACAGAGCATCACCTTTCAGTATTTAAGAGCTTCAAAAGAGTCAGATTAAC	195
QY	315	TTGAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCACTGCATTAAGTGAATTTCTG	374
Db	196	TTGAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCACTGCATTAAGTGAATTTCTG	255
QY	375	GGAAAGGAATGAAGTATATTTTGGTCACAGCTTACATACATAGGAAGTCACACCTGGCT	434
Db	256	GGAAAGGAATGAAGTATATTTTGGTCACAGCTTACATACATAGGAAGTCACACCTGGCT	315
QY	435	CCGCGCAATCCAGACAGAGCTTCTCATCTCCCTCCGCGCTTCTCGCAGTGGATGG	494
Db	316	CCGCGCAATCCAGACAGAGCTTCTCATCTCCCTCCGCGCTTCTCGCAGTGGATGG	375
QY	495	AAACAGTGAAGATGGACCCCACTCAATTAATCTATGATCTCTTATATGGCATCTCCAAG	554
Db	376	AAACAGTGAAGATGGACCCCACTCAATTAATCTATGATCTCTTATATGGCATCTCCAAG	435
QY	555	CTGGGCGCCAGGCAAAATATGAATTTGAGCGCAGCATGACACCACTCCAGGCTGTG	614

D	436	CTGGGGCCAGGGGAAAAAGTATGAAATTGGACGACGACTGACACCACTCCGACCGGTGGTG	495
Q	615	GTCCATGTATGTGAGAGTATTCAGAGAAAGGAGAAACAAGAAATTGGAAAGAAATGAGG	674
D	496	GTCATGTATGTGAGAGTGATCAAGAGAGGAGGAAAGAAATGAAAGATGAGG	555
Q	675	AGACCTTAAGCCAAAAATTATCCAGACCCAGAGCCGAGTACACGCGATCCACCTCAAC	734
D	556	AGACCTTAAGCCAAAAATTATCCAGACCCAGAGCCGAGTACACGCGATCCACCTCAAC	615
Q	735	TGAATCTGGACCGGACGAGAGCGCATTCCTCAATCATACTCACGGAGGAATCTTTTAACTG	794
D	616	TGAATCTGGACCGGACGAGAGCGCATTCCTCAATCATACTCACGGAGGAATCTTTTAACTG	675
Q	795	TGAGAGGTGGTGTCAAGACTCTTCGAGGGTGGAGCCAGATTCGGGGTGGCAGAAATC	854
D	676	TGAGAGGTGGTGTCAAGACTCTTCGAGGGTGGAGCCAGATTCGGGGTGGCAGAAATC	735
Q	855	CCAGTTCATGTGCTCAGAGAGAAATCAAGGCGGTGTCCCTTGTCTTAAATGCTGCACAC	914
D	736	CCAGTTCATGTGCTCAGAGAGAAATCAAGGCGGTGTCCCTTGTCTTAAATGCTGCACAC	795
Q	915	CAGTTACTGTTCATGCGACCCCGGAAATGACTTGGGCGAAATCAGAGTTGTGGTATGAC	974
D	796	CAGTTACTGTTCATGCGACCCCGGAAATGACTTGGGCGAAATCAGAGTTGTGGTATGAC	855
Q	975	CACAAGACATTTGGGAGCTGTCTTTGAGAAACAGATTAATGATGTGTTTGAATCTTGTTC	1034
D	856	CACAAGACATTTGGGAGCTGTCTTTGAGAAACAGATTAATGATGTGTTTGAATCTTGTTC	915
Q	1035	TTTTCTGGTAAAGTTCTGTCTGTGTCCAAAGGCGAGTTGATTCAGTGAAGTCAGGAGAGAGCT	1094
D	916	TTTTCTGGTAAAGTTCTGTCTGTGTCCAAAGGCGAGTTGATTCAGTGAAGTCAGGAGAGAGCT	975
Q	1095	TCCGTTTCTTAAGTGGCCTGCAGGGCCCACTCTACTGTGTAGGAAAGGTATCCACAGAA	1154
D	976	TCCGTTTCTTAAGTGGCCTGCAGGGCCCACTCTACTGTGTAGGAAAGGTATCCACAGAA	1035
Q	1155	AGCCGCTTAAGTCAGAGAGGTTGTGAAAACAGACGAATGCATGTGGAAATTGTAGCGT	1214
D	1036	AGCCGCTTAAGTCAGAGAGGTTGTGAAAACAGACGAATGCATGTGGAAATTGTAGCGT	1095
Q	1215	TTCCTTTCTCCCATGTTCTCATGTTGTGTGATGATATTACTAGATTTCACAGACTAA	1274
D	1096	TTCCTTTCTCCCATGTTCTCATGTTGTGTGATGATATTACTAGATTTCACAGACTAA	1155
Q	1275	CCCTTGTCTGATATTAAGTTACACCGTTGTGTTTATCATCTTTTGGGAAACGAGAAA	1334
D	1156	CCCTTGTCTGATATTAAGTTACACCGTTGTGTTTATCATCTTTTGGGAAACGAGAAA	1215
Q	1335	GGGTTTGGAAAAAGTATCACCTTCCCAAGTCTCGGATTCCTGAGACTCTTTTGCAACAGCA	1394
D	1216	GGGTTTGGAAAAAGTATCACCTTCCCAAGTCTCGGATTCCTGAGACTCTTTTGCAACAGCA	1275
Q	1395	CTTGCTTGGGGAACCTTCTCGAATGCATTCAGCATCCGAAACCGTGACACGCT	1454
D	1276	CTTGCTTGGGGAACCTTCTCGAATGCATTCAGCATCCGAAACCGTGACACGCT	1335
Q	1455	AACCTGTGCTTTTGCAAAAGATTAATCTGAAATTCCTCTGTAAATTTAGCTAATCA	1514
D	1336	AACCTGTGCTTTTGCAAAAGATTAATCTGAAATTCCTCTGTAAATTTAGCTAATCA	1395
Q	1515	ATTACAGAAATGACGTTTCACTGCCAACTTTAGTGGGTGAGAAATTTAGTTAAGTG	1574
D	1396	ATTACAGAAATGACGTTTCACTGCCAACTTTAGTGGGTGAGAAATTTAGTTAAGTG	1455
Q	1575	TTTGGATCGACCTCAGTTCTGTGTTCTTTATGTGTGTCTTATCATGAAATC	1634
D	1456	TTTGGATCGACCTCAGTTCTGTGTTCTTTATGTGTGTCTTATCATGAAATC	1515
Q	1635	ATAGCCAAAACCTTTTGGAAATCTGTTGTGTGAGATGTGTTCTTTACCCACGAA	1694



QY 615 GTCCATGATGTGAGATGATCAAGAGAAAGAGAGAAAGAGAAATGGAATGAG 674  
 Db 496 GTCCATGATGTGAGATGATCAAGAGAAAGAGAGAAAGAGAAATGGAATGAG 555  
 QY 675 AGACCTTAAGCCAAAATTAATTCAGACCAAGAGAGCCGAGTACAGCCGATCCACTGAGC 734  
 Db 556 AGACCTTAAGCCAAAATTAATTCAGACCAAGAGAGCCGAGTACAGCCGATCCACTGAGC 615  
 QY 725 TGAATGAGCAGCAGAGAGAGAGATTCGAATCACTACAGGAGAGATCTTTTACTG 794  
 Db 616 TGAATGAGCAGCAGAGAGAGAGATTCGAATCACTACAGGAGAGATCTTTTACTG 675  
 QY 795 TGAAGTGTGCTGATCAGCACTCTTCGAGAGTGGAGAGAGAGAGAGAGAGAGAG 854  
 Db 676 TGAAGTGTGCTGATCAGCACTCTTCGAGAGTGGAGAGAGAGAGAGAGAGAGAG 735  
 QY 855 CCAATTCATGTTGCTCAGAGAGAGATCAAGAGCCGATCCCTTGTCTTAATGCTGAGC 914  
 Db 736 CCAATTCATGTTGCTCAGAGAGAGATCAAGAGCCGATCCCTTGTCTTAATGCTGAGC 795  
 QY 915 CAGTTACTGTTACAGCAGCCGAGAGATCTTGGCCAACTACAGTGTGTGTGATG 974  
 Db 796 CAGTTACTGTTACAGCAGCCGAGAGATCTTGGCCAACTACAGTGTGTGTGATG 855  
 QY 975 CACAGGACATTTGGAGACTGTCTTGAGAAAAGATATGATAGTGTGTGTGTGATG 1034  
 Db 856 CACAGGACATTTGGAGACTGTCTTGAGAAAAGATATGATAGTGTGTGTGTGATG 915  
 QY 1035 TTTTCTGATGATGCTGT 1094  
 Db 916 TTTTCTGATGATGCTGT 975  
 QY 1095 TCTGT 1154  
 Db 976 TCTGT 1035  
 QY 1155 AGCCGCTATGAGAGAGAGTGTGAAAACAGCAGCAATCATGTGAAATTTGAGCGT 1214  
 Db 1036 AGCCGCTATGAGAGAGAGTGTGAAAACAGCAGCAATCATGTGAAATTTGAGCGT 1095  
 QY 1215 TTTCTTCTTCCCTCATGTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1274  
 Db 1096 TTTCTTCTTCCCTCATGTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1155  
 QY 1275 CTTTGT 1334  
 Db 1156 CTTTGT 1215  
 QY 1335 GCGTTTGGAAAAGATACCTTCCAGATTCGAGATTCGACTCTTTGCAACAGCA 1394  
 Db 1216 GCGTTTGGAAAAGATACCTTCCAGATTCGAGATTCGACTCTTTGCAACAGCA 1275  
 QY 1395 CTTTGT 1454  
 Db 1276 CTTTGT 1335  
 QY 1455 AACTTGT 1514  
 Db 1336 AACTTGT 1395  
 QY 1515 ATTGAGAGATAGCAGTTTCACTGCACTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1574  
 Db 1396 ATTGAGAGATAGCAGTTTCACTGCACTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1455  
 QY 1575 TTTGGAGATCGAAGCTCAGTTTCTGT 1634  
 Db 1456 TTTGGAGATCGAAGCTCAGTTTCTGT 1515  
 QY 1635 ATAGCAAAAATTTTGT 1694  
 Db 1516 ATAGCAAAAATTTTGT 1575

QY 1695 GACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATTATACCTGTGTACCTT 1754  
 Db 1576 GACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATTATACCTGTGTACCTT 1635  
 QY 1755 GGGTGAAGAGTATGAG 1814  
 Db 1636 GGGTGAAGAGTATGAG 1695  
 QY 1815 GTAGAAAAGCTGTCTCCATGCTTAAACAGAGACTGAAATCAAGAGAGTCAAGTGTATG 1874  
 Db 1696 GTAGAAAAGCTGTCTCCATGCTTAAACAGAGACTGAAATCAAGAGAGTCAAGTGTATG 1755  
 QY 1875 TAGATCTGTGCGATATGAG 1934  
 Db 1756 TAGATCTGTGCGATATGAG 1815  
 QY 1935 GTTATCTAGGTTAACTGTTGTAATTTTTTCTAGTTGTATGTGTGTGTGTGTGTGTGTGTGTGT 1994  
 Db 1816 GTTATCTAGGTTAACTGTTGTAATTTTTTCTAGTTGTATGTGTGTGTGTGTGTGTGTGTGTGT 1875  
 QY 1995 GGTATTAATTTTGGCCCTTAAACATACCGTAAACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2054  
 Db 1876 GGTATTAATTTTGGCCCTTAAACATACCGTAAACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1935  
 QY 2055 CCAAGTAAACAATGATGCTTTGGAAATTTGGAAATGCTTTTATCTTGGAGAGCAAT 2114  
 Db 1936 CCAAGTAAACAATGATGCTTTGGAAATTTGGAAATGCTTTTATCTTGGAGAGCAAT 1995  
 QY 2115 ATGTTTGATTAATTTGCTTGTATGTTGTTATATCAATCAAGAAATGTAATGTAATGCTTCAAAAC 2174  
 Db 1996 ATGTTTGATTAATTTGCTTGTATGTTGTTATATCAATCAAGAAATGTAATGTAATGCTTCAAAAC 2055  
 QY 2175 CTGTTTACGTAATCTTGTATGAG 2234  
 Db 2056 CTGTTTACGTAATCTTGTATGAG 2115  
 QY 2235 AGT 2294  
 Db 2116 AGT 2175  
 QY 2295 TGCATCTATTAATTAATGAGAGAGAGATGCTCTTAC 2331  
 Db 2176 TGCATCTATTAATTAATGAGAGAGATGCTCTTAC 2212  
 RESULT 6  
 BC002864 2289 bp mRNA linear PRI 12-JUL-2001  
 LOCUS BC002864  
 DEFINITION Homo sapiens, Down syndrome critical region gene 1, clone MGC:10263  
 IMAGE:3944959, mRNA, complete cde.  
 ACCESSION BC002864  
 VERSION BC002864.1 GI:12804022  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2289)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbhe@mail.nih.gov](mailto:cgapbhe@mail.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arriayed by: The I.M.A.G.E. Consortium (ILLUM)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>  
Contact: nisc\_mcgcngr@nih.gov  
Shevchenko, Y., Wetherby, K.D., Beckerrom-  
Benjamin, B., Blakesley, R.W., Beckerrom-  
Dietrich, N., Guan, X., Gupta, J., Ho, S.,  
Lim, M., Maduro, C.L., Mastello, C., Mastri-  
McDowell, J., Pearson, R., Snyder, B., Stan-  
tingson, E.E., Touchman, J.W., Taube, C.  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 15 Row: a Column: 23  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 7596913.

source

## Source

CD5

```

CDS
        /name="vector: pDHB/"
        /size=66..659
        /codon_start=1
        /product="Down syndrome critical region gene 1"
        /protein_id="AAH03864.1"
        /db_xref="GI:12804023"
        /translation="MEEVDLDPSSATIACHLDPRFVVDGGLCRAPFEC
        FOYPKSPKRVINFSNPSADARIOLKHTEPLGEMKLYPAQTHLHPC
        KQFLSPASPPVGVQVEDATVPYDLYLVAISKLGCEKTELHAAV
        CESDQEKESMEWRAPKPKIITRAPETTPHLG"
        628 a 465 c 555 g 641 t
        BASE COUNT
        ORIGIN

```

Query Match	90.4%	Score 2107.4	DB 9	Length 2289				
Best Local Similarity	97.5%	Pred. No. 0						
Matches 2141	Conservative 0	Mismatches 56	Indels 0	Gaps 0				
135	GAAGCAGATGCATTTT	TAGAACTTT	AAC	TAC	CAGTTT	TAGTCCCTGATTC	CGCTGTGTG	194
Db								
57	GACTGCGAGATGGAGG	AGGTGGACCT	CGACGACCT	CGCCAGCGCCAC	CACCTGC	CGCTGTAC	116	
195	GCAACACGTGATATCT	TTCAGCGAAAGTGA	AAACAGAGGGCCAA	ATTTGAGTCCCTCT	TTTAGG	254		
Db								
117	CTGACCCGCGGTGTC	TGCTGGACGGCTG	CGCGGGCCAA	ATTTGAGTCCCTCT	TTTAGG	176		
255	ACGTATGACAGGACAT	CACTTTCAGTATTT	TAAAGAGCTTCA	AAACGAGTCA	GAAATAAC	314		
Db								
177	ACGTATGACAGGACAT	CACTTTCAGTATTT	TAAAGAGCTTCA	AAACGAGTCA	GAAATAAC	236		
315	TTCAGCAACCCCTTC	CGCGAGCAGATGC	CAGGCTCCAGTGC	ATAAGATG	TTTCTG	374		
Db								
237	TTCAGCAACCCCTTC	CGCGAGCAGATGC	CAGGCTCCAGTGC	ATAAGATG	TTTCTG	296		
375	GGAAAGGAAATGAAG	TTTATATTTTGTCT	CAGACTTACATAGG	AGCTCACCTGGCT	434			
Db								
297	GGAAAGGAAATGAAG	TTTATATTTTGTCT	CAGACTTACATAGG	AGCTCACCTGGCT	356			
435	CCGCCAAATCCAGAC	AGAGTTTCTGATCT	CCCTCCCGCTCT	CGCCAGTGGGATGG	494			
Db								
357	CCGCCAAATCCAGAC	AGAGTTTCTGATCT	CCCTCCCGCTCT	CGCCAGTGGGATGG	416			
495	AAACAAGTGAAGATGC	ACCCAGTCATAA	CTATGATCTCTTAT	ATGCCATCTCCAAG	554			
Db								
417	AAACAAGTGAAGATGC	ACCCAGTCATAA	CTATGATCTCTTAT	ATGCCATCTCCAAG	476			
555	CTGGGCGCAGGGGAA	AGTATGAATTTG	CACGAGCGATG	CACCACTCCAGCGTGGT	614			
Db								
477	CTGGGCGCAGGGGAA	AGTATGAATTTG	CACGAGCGATG	CACCACTCCAGCGTGGT	536			

```

Db 1617 GACATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCATCTGTTGTAACCTT 1676
Oy 1755 GGGTGAAGAATGAGCAGTGGGAGACTAAGATGATTAACCTACCTGTAATCATATGTT 1814
Db 1677 GGGTGAAGAATGAGCAGTGGGAGACTAAGATGATTAACCTACCTGTAATCATATGTT 1736
Oy 1815 GTAGAAAAAGCTGTTCCCTGCTTAACAGAGCTTGAATTCAAAGCATGTCAAGTGATAG 1874
Db 1737 GTAGAAAAAGCTGTTCCCTGCTTAACAGAGCTTGAATTCAAAGCATGTCAAGTGATAG 1796
Oy 1875 TAGATCTGTGGGATATGAGAGGATGAGTGCCTTCCCTCCATTCCTGATGGAATTT 1934
Db 1797 TAGATCTGTGGGATATGAGAGGATGAGTGCCTTCCCTCCATTCCTGATGGAATTT 1856
Oy 1935 GTTATATAGGTTAACTGTTGTAATTTTTTTTCTAGTTGTAATGCTATGCTGTAATA 1994
Db 1857 GTTATATAGGTTAACTGTTGTAATTTTTTTTCTAGTTGTAATGCTATGCTGTAATA 1916
Oy 1995 GGTATATATTTTGGGCTTACAAATCCGTAACAGTGTTCATTTTGAAATCTTAATG 2054
Db 1917 GGTATATATTTTGGGCTTACAAATCCGTAACAGTGTTCATTTTGAAATCTTAATG 1976
Oy 2055 CCAGTATACATGATGCTGTTGGAATTTGGAAGATGCTTTATTTCTTGAGAGCAAT 2114
Db 1977 CCAGTATACATGATGCTGTTGGAATTTGGAAGATGCTTTATTTCTTGAGAGCAAT 2036
Oy 2115 ATGTTTGCATTAATGCTTTGATTTGTCATATCAAGAAATTTGATGGAAGTCTCAACC 2174
Db 2037 ATGTTTGCATTAATGCTTTGATTTGTCATATCAAGAAATTTGATGGAAGTCTCAACC 2096
Oy 2175 CTGTTTACGCTACTTGTGTAAGAGGAGCCGTTTGGAGAGACCATTTGATGCTGTCCA 2234
Db 2097 CTGTTTACGCTACTTGTGTAAGAGGAGCCGTTTGGAGAGACCATTTGATGCTGTCCA 2156
Oy 2235 AGTGTTCCTTGTAGTGTCTTTTAACTGAGAGGCTAACCTCAAAATCTTTTAACT 2294
Db 2157 AGTGTTCCTTGTAGTGTCTTTTAACTGAGAGGCTAACCTCAAAATCTTTTAACT 2216
Oy 2295 TGCATTCATTAATTAATGGGACAGATAGCTCCTTAC 2331
Db 2217 TGCATTCATTAATTAATGGGACAGATAGCTCCTTAC 2253

RESULT 7
AK092184 2407 bp mRNA linear PRI 15-JUL-2002
AK092184 Homo sapiens cDNA FLJ34865 f18, clone NT2NE2014104, highly similar
LOCUS AK092184
DEFINITION AK092184.1 GI:21750714
ACCESSION AK092184
VERSION AK092184.1
KEYWORDS oligo capping; f18 (full insert sequence);
SOURCE Homo sapiens testocarcinoma cell line:NT2 cDNA to mRNA,
clone lib:NT2NE2 clone:NT2NE2014104.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Nishimura, K., Wagaruma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omiya, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kikuchi, H., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Oshima, A., Sugiyama, A., Kaneko, K., Takahashi, Fujii, A.,
Negahari, K., Maehara, Y., Negai, K. and Isegai, T.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 2407)
Isegai, T. and Yamamoto, J.
Direct Submission
JOURNAL
TITLE
AUTHORS
SUBMITTED (04-JUL-2002) Takao Isegai, FLJ Project (HRI Team); 2-6-7

```

```

COMMENT Kazuo-Kamatari, Kiearazu, Chiba 292-0812, Japan
(E-mail:genom@shri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'-83'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES
source
1..2407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2NE2014104"
/cell_line="NT2"
/cell_type="testocarcinoma"
/clone_lib="NT2NE2"
/notes="Cloning vector: pMB18SFL3-mRNA from NT2 neuron
after the differentiation of NT2 neuronal precursor
cells."

BASE COUNT 607 a 519 c 627 g 654 t

ORIGIN
Query Match 90.3%; Score 2105.8; DB 9; Length 2407;
Best Local Similarity 97.44; Pred. No. 0;
Matches 2140; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Oy 135 GAAAGCAAGATGATTTTAAAGAACTTTAACTACAGTTTAACTCCCTGATTTGCTGTGTG 194
Db 209 GATGCGAGATGAGAGAGGTGAGACCTGACAGACTGCGCAGCGCACATGCTGTGTGAC 268
Oy 195 GCAACAGTATATCTTACGAGGAAAGTAAACAGGCGCAATTTAGTCCCTTTAGG 254
Db 269 CTGAGCCCGCGGTGCTGAGAGGCGCTGTGCGGCGCAATTTGATGCTCTTTAGG 328
Oy 255 ACGTATGCAAGGACATACCTTTCAGTATTTTAAAGCTTCAACAGAGTCAGATTAAC 314
Db 329 ACGTATGCAAGGACATACCTTTCAGTATTTTAAAGCTTCAACAGAGTCAGATTAAC 388
Oy 315 TTCAGCAACCCCTTCTCCGAGCAGATGCGAGCTCCAGCTGCAATTAAGCTGATTTCTG 374
Db 369 TTCAGCAACCCCTTCTCCGAGCAGATGCGAGCTCCAGCTGCAATTAAGCTGATTTCTG 448
Oy 375 GGAAGGAAATGAAGTATTTTGTCTCAGACCTTACCATAGGAGCTCACACTGGCT 434
Db 449 GGAAGGAAATGAAGTATTTTGTCTCAGACCTTACCATAGGAGCTCACACTGGCT 508
Oy 435 CCGCAAAATCCAGACAGACAGTTCTGATCTCCCTCCGCTCTCCGCAAGTGGATGG 494
Db 509 CCGCAAAATCCAGACAGACAGTTCTGATCTCCCTCCGCTCTCCGCAAGTGGATGG 568
Oy 495 AAACAAATGGAAGATGCGACCCCACTCAATTAATGATTTCTTATATGCAATCTCCAA 554
Db 569 AAACAAATGGAAGATGCGACCCCACTCAATTAATGATTTCTTATATGCAATCTCCAA 628
Oy 555 CTGGGCGCAGAGGAAAGATGAATTTGACGCGAGGAGCTGACACCACTCCAGCGTGTG 614
Db 629 CTGGGCGCAGAGGAAAGATGAATTTGACGCGAGGAGCTGACACCACTCCAGCGTGTG 688
Oy 615 GTTCATGTATGTGAGAGTATCAAGAGAGAGAAAGAAAGAAAGAAATGAAGAG 674
Db 689 GTTCATGTATGTGAGAGTATCAAGAGAGAGAGAAAGAAAGAAATGAAGAG 748
Oy 675 AGACCTAACCCAAATTTATTCAGACCAAGAGGCGGAGATACACCCCATCCACTCGC 734
Db 749 AGACCTAACCCAAATTTATTCAGACCAAGAGGCGGAGATACACCCCATCCACTCGC 808
Oy 735 TGAATGCGACGCGAGAGAGAGCATTTCCAAATCATATCAAGAGAGAAATCTTTACTG 794
Db 809 TGAATGCGACGCGAGAGAGAGCATTTCCAAATCATATCAAGAGAGAAATCTTTACTG 868
Oy 795 TGAGAGTGTGTGTCAGACTTCTTGGAGGTGGACCGGAGATCGGGGTGCGAATAATC 854

```



Db 869 TGAGGTGGCTGTGTCAGACCTCTTCTCGAGGTGGCAGCGAGATCGGGGTGCAGAAATC 928  
QY 855 CCAAGTTATGTTCTCAGAGAGAAATCAAGGGCGTGTCCCTTGTCTTAATGCTGCACAC 914  
Db 929 CCAAGTTATGTTCTCAGAGAGAAATCAAGGGCGTGTCCCTTGTCTTAATGCTGCACAC 988  
QY 915 CAGTTACTGTTCTATGCGACCCGGGAATGACTTTGGGCCAATCACTAGTTTGTGGTGATCG 974  
Db 989 CAGTTACTGTTCTATGCGACCCGGGAATGACTTTGGGCCAATCACTAGTTTGTGGTGATCG 1048  
QY 975 CACAAGCAATTTGGGACTGTTCTTGAGAAACAGATAATGATAGTTTGTGTAATGTTTC 1034  
Db 1049 CACAAGCAATTTGGGACTGTTCTTGAGAAACAGATAATGATAGTTTGTGTAATGTTTC 1108  
QY 1035 TTTTCTGTTAGTGTCTGTTCTGTCGCAAGGGCAGGTTGATCAGTGAAGCTCAGCAGAGAGCT 1094  
Db 1109 TTTTCTGTTAGTGTCTGTTCTGTCGCAAGGGCAGGTTGATCAGTGAAGCTCAGCAGAGAGCT 1168  
QY 1095 TCCTGTTTCTAAGTGTGCTGAGGGGCCACTCTCTACTGTTAGGAAGGTACCACAGGA 1154  
Db 1169 TCCTGTTTCTAAGTGTGCTGAGGGGCCACTCTCTACTGTTAGGAAGGTACCACAGGA 1228  
QY 1155 AGCCGCTAGTGCAGAGAGTTTGTGAAAACAGCAGCAGCAATGCAATGTGCAATTTGACGTT 1214  
Db 1229 AGCCGCTAGTGCAGAGAGTTTGTGAAAACAGCAGCAGCAATGCAATGTGCAATTTGACGTT 1288  
QY 1215 TTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGCAATGATATTTACTGATTTTACAAGACTAA 1274  
Db 1289 TTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGCAATGATATTTACTGATTTTACAAGACTAA 1348  
QY 1275 CCTTTGTTCTGATATAAAGTTTACACCGTTGTTGTTTACATCTTTTGGGAAGCCAGAAA 1334  
Db 1349 CCTTTGTTCTGATATAAAGTTTACACCGTTGTTGTTTACATCTTTTGGGAAGCCAGAAA 1408  
QY 1335 GCGTTTGGAAAACGTATCACCTTTCCAGATTTCTCGAATTTCTCGACTTTTGGCAACAGCA 1394  
Db 1409 GCGTTTGGAAAACGTATCACCTTTCCAGATTTCTCGAATTTCTCGACTTTTGGCAACAGCA 1468  
QY 1395 CTTCGTTGCGGAACCTCTTCGGAAGTCATCTCAGCATCCCCCAACCGTGCAACGTGT 1454  
Db 1469 CTTCGTTGCGGAACCTCTTCGGAAGTCATCTCAGCATCCCCCAACCGTGCAACGTGT 1528  
QY 1455 AACTGTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGAGAAATTTAGCTTATACA 1514  
Db 1529 AACTGTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGAGAAATTTAGCTTATACA 1588  
QY 1515 ATTACAGAAATACAGTTTCACTGCCAACTTTTAGTGGGTGAGAAATTTTAGTTAGGTG 1574  
Db 1589 ATTACAGAAATACAGTTTCACTGCCAACTTTTAGTGGGTGAGAAATTTTAGTTAGGTG 1648  
QY 1575 TTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGTCTATACATGAATC 1634  
Db 1649 TTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGTCTATACATGAATC 1708  
QY 1635 ATAGCCAAAACCTTTTGGAAACTGTTGGTTGAGATAGTGGTCTTTTACCCACAGAA 1694  
Db 1709 ATAGCCAAAACCTTTTGGAAACTGTTGGTTGAGATAGTGGTCTTTTACCCACAGAA 1768  
QY 1695 GACATCAAGATACACTTGTAAATTAAGCTAGATGATATATTATCATCTGTTGTACACTT 1754  
Db 1769 GACATCAAGATACACTTGTAAATTAAGCTAGATGATATATTATCATCTGTTGTACACTT 1828  
QY 1755 GGGTGAAGATAGTGGCAGTGGGAGACTAGATGATATTAACTTACCTTACCTGTGAATCATGTT 1814  
Db 1829 GGGTGAAGATAGTGGCAGTGGGAGACTAGATGATATTAACTTACCTTACCTGTGAATCATGTT 1888  
QY 1815 GTAGGAAAGCTGTTCCCATGTTCAACAGGACTTGAATTTCAAGCATGTTCAAGTGGATAG 1874  
Db 1889 GTAGGAAAGCTGTTCCCATGTTCAACAGGACTTGAATTTCAAGCATGTTCAAGTGGATAG 1948  
QY 1875 TAGATCTGTGGCGATATGAGAGGATGAGTGCCTTTCCCATTCATCTTCCTGATGGAATT 1934

Db 1949 TAGATCTGTGGCGATATGAGAGGATGACGTGCTTTTCCCATTCATCTCTGATGCAATT 2008  
QY 1935 GTTATACCTAGGTTAAACATTTTGTAAATTTTCTAGTTGTAATGCTGATGCTGGTAAATA 1994  
Db 2009 GTTATACCTAGGTTAAACATTTTGTAAATTTTCTAGTTGTAATGCTGATGCTGGTAAATA 2068  
QY 1995 GGTATTATATTTTGGGCTTACAATACCGTAACAATGTTTGTTCATTTTGAATACTTAAATG 2054  
Db 2069 GGTATTATATTTTGGGCTTACAATACCGTAACAATGTTTGTTCATTTTGAATACTTAAATG 2128  
QY 2055 CCAAGTAAACAATGCAATGCTTTTGGAAATTTTGGAAATGTTTATTTTGAAGACCAAT 2114  
Db 2129 CCAAGTAAACAATGCAATGCTTTTGGAAATTTTGGAAATGTTTATTTTGAAGACCAAT 2188  
QY 2115 ATGTTTGCATTAATGCTTTTGAATGTTTGCATATCAAGAATTTGATTAACCTTCTCAAC 2174  
Db 2189 ATGTTTGCATTAATGCTTTTGAATGTTTGCATATCAAGAATTTGATTAACCTTCTCAAC 2248  
QY 2175 CTGTTTACCGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCATTCATCGCTGTCCA 2234  
Db 2249 CTGTTTACCGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCATTCATCGCTGTCCA 2308  
QY 2235 AGTGTTCCTTGTAAAGTCTTTTAACTGGAGAGCTAACCTCABAATACTTTTTTTAAAC 2294  
Db 2309 AGTGTTCCTTGTAAAGTCTTTTAACTGGAGAGCTAACCTCABAATACTTTTTTTAAAC 2368  
QY 2295 TGCATCTTAATAAATGGGCGACAGTATGCTCCTTAC 2331  
Db 2369 TGCATCTTAATAAATGGGCGACAGTATGCTCCTTAC 2405

RESULT 8  
LOCUS HSU85265  
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,  
alternative exon 1, complete cds.  
ACCESSION U85265  
VERSION U85265.2 GI:7596912  
KEYWORDS Homo sapiens.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2284)  
AUTHORS Fuentetaja, J., Pritchard, M.A., Planas, A.M., Bosch, A., Ferrer, I. and Estivill, X.  
TITLE A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart  
JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)  
MEDLINE 96121593  
PUBMED 8595418  
REFERENCE 2 (bases 1 to 2284)  
AUTHORS Fuentetaja, J., Pritchard, M.A. and Estivill, X.  
TITLE Genomic organization, alternative splicing, and expression patterns of the DSCR1 (Down syndrome candidate region 1) gene  
JOURNAL Genomics 44 (3), 358-361 (1997)  
MEDLINE 97468152  
PUBMED 9325060  
REFERENCE 3 (bases 1 to 2284)  
AUTHORS Fuentetaja, J., Pritchard, M. and Estivill, X.  
TITLE Direct Submission  
JOURNAL Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain  
REFERENCE 4 (bases 1 to 2284)  
AUTHORS Fuentetaja, J., Pritchard, M. and Estivill, X.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain  
REMARK Sequence update by submitter  
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612864.  
FEATURES Location/Qualifiers

source	1..2284	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="21"	
		/map="21q22.1-q22.2"	
gene	1..2284		
		/gene="DSCR1"	
	322..675	/gene="DSCR1"	
CDS		/note="down syndrome candidate region 1; one of four alternatively spliced exon 1"	
		/codon_start=1	
		/protein_id="AA084370.1"	
		/db_xref="GI:2618743"	
		/translation="MKLYFAQTLHIGSSHLAPNPDKFLISPPAPVKKOVEDATPVINDLYAISKLGPGEKELHATDTPSVVHVCSDEKSEEMERKPKKIIOTRRPEYPIHLS"	
BASE COUNT	626 a 468 c 534 g 656 t		
ORIGIN			
Query Match	90.3%; Score 2104; DB 9; Length 2284;		
Best Local Similarity	99.1%; Pred. No. 0;		
Matches 2116; Conservative	0; Mismatches 20; Indels 0; Gaps 0;		
QY	196	CAACAGTATCTTCAGGAAAGTAAACCGAGGCAAAATTTGATCCCTTTAGGA	255
DB	134	CAATCTGACATATGGAAAGAAAGCCAAATTTGGCTCTCTTTAGGA	193
QY	256	CGATGACAGACATCACTTCAGTATTTTAAAGCTTCAACGAGTCAGATAACT	315
DB	194	CGATGACAGACATCACTTCAGTATTTTAAAGCTTCAACGAGTCAGATAACT	253
QY	316	TCAGCAACCCCTTCTCCGACAGAGTCCAGGCTCCAGCTGCAATAGAGCTTCTG	375
DB	254	TCAGCAACCCCTTCTCCGACAGAGTCCAGGCTCCAGCTGCAATAGAGCTTCTG	313
QY	376	GAAGGAAATGAAGTTATTTTGTCTCAGACCTTACACATAGGAAGTCAACCTGCTC	435
DB	314	GAAGGAAATGAAGTTATTTTGTCTCAGACCTTACACATAGGAAGTCAACCTGCTC	373
QY	436	CGCAAAATCCAGACAGAGTTCTGATCTCCCTCCCTCTCCGCGCAGTGGATGGA	495
DB	374	CGCAAAATCCAGACAGAGTTCTGATCTCCCTCCCTCTCCGCGCAGTGGATGGA	433
QY	496	AACAATGGAAGATCCAGCCCACTCAATTAATGATCTCTTATATGCACTCCAGC	555
DB	434	AACAATGGAAGATCCAGCCCACTCAATTAATGATCTCTTATATGCACTCCAGC	493
QY	556	TGGGGCCAGGAAAGATGATGATTCGACGAGGACCTGACACCACTCCAGCTGCTG	615
DB	494	TGGGGCCAGGAAAGATGATGATTCGACGAGGACCTGACACCACTCCAGCTGCTG	553
QY	616	TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	675
DB	554	TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	613
QY	676	GACCTAAGCCAAATTTATCCAGACGAGAGGCGGAGTACACCGGATCCACCTCACT	735
DB	614	GACCTAAGCCAAATTTATCCAGACGAGAGGCGGAGTACACCGGATCCACCTCACT	673
QY	736	GAATGGAACGCGAGAGAGAGATTCGAATCATCTCAAGGAGGAATCTTTTACTGT	795
DB	674	GAATGGAACGCGAGAGAGAGATTCGAATCATCTCAAGGAGGAATCTTTTACTGT	733
QY	796	GAGAGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	855
DB	734	GAGAGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	793
QY	856	CAGTTATGTTGCTCAGAGAGATCAAGCCGTGCTCCCTTTTCTAATGCTGACACC	915
DB	794	CAGTTATGTTGCTCAGAGAGATCAAGCCGTGCTCCCTTTTCTAATGCTGACACC	853
QY	916	AGTTACTGTTTATGACACCGGAAATGACTTGGCCAAATCACTGAGTTTGTGTATGCG	975
DB	854	AGTTACTGTTTATGACACCGGAAATGACTTGGCCAAATCACTGAGTTTGTGTATGCG	913
QY	976	ACAGAGACATTTGGAGCTCTCTTGAAGAAACAGATATGATGATTTTGTACTGTTCT	1035
DB	914	ACAGAGACATTTGGAGCTCTCTTGAAGAAACAGATATGATGATTTTGTACTGTTCT	973
QY	1036	TTTCTGTAGGTTCTGTCTGTGCAAGGAGGAGGAGTGTATCAGTGAAGCTCAGAGAGCTT	1095
DB	974	TTTCTGTAGGTTCTGTCTGTGCAAGGAGGAGGAGTGTATCAGTGAAGCTCAGAGAGCTT	1033
QY	1096	CTGTGTTCTAAGTGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1155
DB	1034	CTGTGTTCTAAGTGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1093
QY	1156	GCGCCCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1215
DB	1094	GCGCCCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1153
QY	1216	TCCTTCTTCCCTCATGTTCTCATGTTTGTGATGATGATGATGATGATGATGATGATG	1275
DB	1154	TCCTTCTTCCCTCATGTTCTCATGTTTGTGATGATGATGATGATGATGATGATGATG	1213
QY	1276	CTTTGTTCTGATATTAAGTTACACCGGTTGTTTACATCTTTGGAGAGGAGGAGGAGG	1335
DB	1214	CTTTGTTCTGATATTAAGTTACACCGGTTGTTTACATCTTTGGAGAGGAGGAGGAGG	1273
QY	1336	CGTTGGAACGATACCTTCTCCAGATTCGATTCGATCTTTGCAACAGCAC	1395
DB	1274	CGTTGGAACGATACCTTCTCCAGATTCGATTCGATCTTTGCAACAGCAC	1333
QY	1396	TTGCTTGGGAACCTTCTCCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	1455
DB	1334	TTGCTTGGGAACCTTCTCCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	1393
QY	1456	ACTTGTGCTTTTGGCAAAAGAGTGTGATGATGATGATGATGATGATGATGATGATG	1515
DB	1394	ACTTGTGCTTTTGGCAAAAGAGTGTGATGATGATGATGATGATGATGATGATGATG	1453
QY	1516	TTACAGAAATAGCAGTTTCACTGCCAATTTTATGAGGAGGAGGAGGAGGAGGAGGAGG	1575
DB	1454	TTACAGAAATAGCAGTTTCACTGCCAATTTTATGAGGAGGAGGAGGAGGAGGAGGAGG	1513
QY	1576	TTGGGATGGAACCTGATGTTGTTTCTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1635
DB	1514	TTGGGATGGAACCTGATGTTGTTTCTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1573
QY	1636	TAGCCAAATCTTTTGGAAACGTTGTTGATGATGATGATGATGATGATGATGATGATG	1695
DB	1574	TAGCCAAATCTTTTGGAAACGTTGTTGATGATGATGATGATGATGATGATGATGATG	1633
QY	1696	ACATCAAGTATACCTGTTAAATTAAGCTGATGATGATGATGATGATGATGATGATGATG	1755
DB	1634	ACATCAAGTATACCTGTTAAATTAAGCTGATGATGATGATGATGATGATGATGATGATG	1693
QY	1756	GATGAAAGATGAGCAGTGGAGACTAAGATGATTAACCTACCTGTAATCATATGTTG	1815
DB	1694	GATGAAAGATGAGCAGTGGAGACTAAGATGATTAACCTACCTGTAATCATATGTTG	1753
QY	1816	TAGAAAGCTGTTCCATGTTCTAACAGACTTGAATTCAGAGCTGCAAGTGAATGAT	1875
DB	1754	TAGAAAGCTGTTCCATGTTCTAACAGACTTGAATTCAGAGCTGCAAGTGAATGAT	1813
QY	1876	AGATCTGAGGATGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGG	1935
DB	1814	AGATCTGAGGATGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGG	1873
QY	1936	TTATATAGGTTAACTTGTAAATTTTCTAGTTGATGATGATGATGATGATGATGATGATG	1995
DB	1874	TTATATAGGTTAACTTGTAAATTTTCTAGTTGATGATGATGATGATGATGATGATGATG	1933
QY	1996	GTAATATTTTGGCTTAATACGTTAATGATGATGATGATGATGATGATGATGATGATG	2055



Db 1934 GTATTATATTTTGGCCCTTACAATACCGTAACATGTTTGTGTCATTTTGAATACTTAATGCG 1993  
QY 2056 CAAGTAACATGATGCTTTTGGAAATTTGGAAGATGGTTTATTTCTTTGAGAGCAATA 2115  
Db 1994 CAAGTAACATGATGCTTTGGAATTTGGAAGATGGTTTATTTCTTTGAGAGCAATA 2053  
QY 2116 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2175  
Db 2054 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2113  
QY 2176 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2235  
Db 2114 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2173  
QY 2236 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2295  
Db 2174 TGTTTCGATTAATGCTTTGATTTGTTTCATATCAAGAAATTTGATTAAGAGTTCTCAAAACC 2233  
QY 2296 GCAATCTATAATAAATGGGACAGTATGCTCTTAC 2331  
Db 2234 GCAATCTATAATAAATGGGACAGTATGCTCTTAC 2269

RESULT 9  
AX410694  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE

AX410694  
Sequence 3341 from Patent WO0229103.  
AX410694  
AX410694.1 GI:21443399  
human.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1  
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.  
Gene expression profiles in liver cancer  
Patent: WO 0229103-A 3341 11-APR-2002;  
GENE LOGIC INC (US)

FEATURES  
source  
1..2173  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/notes="EMBL/GenBank Accession No. U28833"  
BASE COUNT  
606 a 431 c 503 g 633 t  
ORIGIN

Query Match 90.1%; Score 2101; DB 6; Length 2173;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 GCCAAATTTGAGTCCCTTTAGGACGTATGACAGGACATCACCTTTGAGTATTTTAAAG 290  
Db 58 GCCAAATTTGAGTCCCTTTAGGACGTATGACAGGACATCACCTTTGAGTATTTTAAAG 117  
QY 291 AGCTTCAAACGAGTCAAGATAAATTTAGCAACCCCTTTCTCGAGCAGATGCCAGGCTC 350  
Db 118 AGCTTCAAACGAGTCAAGATAAATTTAGCAACCCCTTTCTCGAGCAGATGCCAGGCTC 177  
QY 351 CAGCTGATTAAGACTGAGTCTTGGGAAAGGAAATGAAGTTATTTTCTCAGACCTTA 410  
Db 178 CAGCTGATTAAGACTGAGTCTTGGGAAAGGAAATGAAGTTATTTTCTCAGACCTTA 237  
QY 411 CACATAGGAGCTCACACCTGGCTCGGCAAAATCAGAGATGGACCCCAAGTCATAAATPAT 470  
Db 238 CACATAGGAGCTCACACCTGGCTCGGCAAAATCAGAGATGGACCCCAAGTCATAAATPAT 297  
QY 471 CCGGCTCTCCGCGAGTGGGATGGAAGTGGAGATGGACCCCAAGTCATAAATPAT 530  
Db 298 CCGGCTCTCCGCGAGTGGGATGGAAGTGGAGATGGACCCCAAGTCATAAATPAT 357  
QY 531 GATCTCTATATGCCATCTCCAAAGCTGGGCGAGGGAAGATGAAATTCAGCGAGCG 590

Db 358 GATCTCTTATATGCCATCTCCAAGCTGGGCGCAGGGAAAAAGTATGAATTTGACGAGCG 417  
QY 591 ACTGACACCACTCCACGCGTGGTCCATGTATCTGAGAGTATCAAGAGAAAGGAGAA 650  
Db 418 ACTGACACCACTCCACGCGTGGTCCATGTATGTGAGAGTATCAAGAGAAAGGAGAA 477  
QY 651 GAAGAGAAATGGAAGAATGAGGAGACCTTAAGCCAAAAATATCCAGACCCAGGAGGCGG 710  
Db 478 GAAGAGAAATGGAAGAATGAGGAGACCTTAAGCCAAAAATATCCAGACCCAGGAGGCGG 537  
QY 711 GAGTACACCCGATCCACCTCAGCTGAACTGGCACCGGACGAGGACGCAATCCAAATCAT 770  
Db 538 GAGTACACCCGATCCACCTCAGCTGAACTGGCACCGGACGAGGACGCAATCCAAATCAT 597  
QY 771 ACTCACGGGAGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 830  
Db 598 ACTCACGGGAGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 657  
QY 831 GCGGAGATCGGGTGGCAGAAATCCAGTTTCTGTGTCAGAGAGAAATCAAGGCCGTG 890  
Db 658 GCGGAGATCGGGTGGCAGAAATCCAGTTTCTGTGTCAGAGAGAAATCAAGGCCGTG 717  
QY 891 TCCCTTTGTTCTAATGCTGCACACCACTTACTGTTTCTGAGTGGTCTGTGTCAGAGGCGAGTT 950  
Db 718 TCCCTTTGTTCTAATGCTGCACACCACTTACTGTTTCTGAGTGGTCTGTGTCAGAGGCGAGTT 777  
QY 951 CAATCACTGAGTTTGTGGTATCGCAAGGACATTTGGGACTGTCTTGAGAAAAACAGAT 1010  
Db 778 CAATCACTGAGTTTGTGGTATCGCAAGGACATTTGGGACTGTCTTGAGAAAAACAGAT 837  
QY 1011 AATGATAGTGTGTTTGTGTTTCTTTTCTGAGTGGTCTGTGTCAGAGGCGAGTT 1070  
Db 838 AATGATAGTGTGTTTGTGTTTCTTTTCTGAGTGGTCTGTGTCAGAGGCGAGTT 897  
QY 1071 GATCAGTGTGCTGAGGAGAGCTTCTGTTTCTAAGTGGCTTCGAGGCGACCTCTCTA 1130  
Db 898 GATCAGTGTGCTGAGGAGAGCTTCTGTTTCTAAGTGGCTTCGAGGCGACCTCTCTA 957  
QY 1131 CTGTTAGGAAAGGTACCAAGGAGCGCTAGTGCAGAGAGTTGTGAAAAACAGCAGC 1190  
Db 958 CTGTTAGGAAAGGTACCAAGGAGCGCTAGTGCAGAGAGTTGTGAAAAACAGCAGC 1017  
QY 1191 AATGCAATGTGGAATTTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1250  
Db 1018 AATGCAATGTGGAATTTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1077  
QY 1251 TATATCTGATTTTCAAGACTAAACCTTTGTTTCTATATAAAGTTACACCGTTGTTCTTT 1310  
Db 1078 TATATCTGATTTTCAAGACTAAACCTTTGTTTCTATATAAAGTTACACCGTTGTTCTTT 1137  
QY 1311 TACATCTTTGGGAAAGCGGAAAGCGTTTGGAAACGTATACCTTTCCAGATTTCTCG 1370  
Db 1138 TACATCTTTGGGAAAGCGGAAAGCGTTTGGAAACGTATACCTTTCCAGATTTCTCG 1197  
QY 1371 GATTTCTGACTCTTTGCAACAGCAGCTTGTGTTGCGAACTCTTCTGGAATGCAATTCAC 1430  
Db 1198 GATTTCTGACTCTTTGCAACAGCAGCTTGTGTTGCGAACTCTTCTGGAATGCAATTCAC 1257  
QY 1431 AGCATCCCAACCGTGCACGTTAACTTTGCTTTTGGAAAGAGTTGATCTGAAAT 1490  
Db 1258 AGCATCCCAACCGTGCACGTTAACTTTGCTTTTGGAAAGAGTTGATCTGAAAT 1317  
QY 1491 CCTCTGTAGAAATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTA 1550  
Db 1318 CCTCTGTAGAAATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTA 1377  
QY 1551 GGGTGAAGAAATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTA 1610  
Db 1378 GGGTGAAGAAATTTAGTTTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTA 1437  
QY 1611 TGTGTTGTTTCTTATACATGAATCATAGCCAAAACTTTTTTGGAAACCTTTGTTGTTGAGA 1670  
Db 1438 TGTGTTGTTTCTTATACATGAATCATAGCCAAAACTTTTTTGGAAACCTTTGTTGTTGAGA 1497

QY 1671 TAGTGGTCTTTTACCCACGAGACATCAAGATACCTGTAAATAAGCTGTAGCA 1730  
 DB 1498 TAGTGGTCTTTTACCCACGAGACATCAAGATACCTGTAAATAAGCTGTAGCA 1557  
 QY 1731 TATATTCATACCTGTTGTACATTTGGGTGAAAAGTATGGAGTGGAGACTAGATGAT 1790  
 DB 1558 TATATTCATACCTGTTGTACATTTGGGTGAAAAGTATGGAGTGGAGACTAGATGAT 1617  
 QY 1791 TAAACCAACCTGTGAATCATATGTGTAGAGAAAAGCTGTGCCATGTCTAACAGACTTGA 1850  
 DB 1618 TAAACCAACCTGTGAATCATATGTGTAGAGAAAAGCTGTGCCATGTCTAACAGACTTGA 1677  
 QY 1851 ATTCAGAGATGCAAGTGAATGATGATCTGTGGGATATATGAGAGGAGTGGAGCCCTT 1910  
 DB 1678 ATTCAGAGATGCAAGTGAATGATGATCTGTGGGATATATGAGAGGAGTGGAGCCCTT 1737  
 QY 1911 TCCCATTCATCCCTGATGGAATGTTATAGTTAGTTAATCATTTGTATTTTCTAGT 1970  
 DB 1738 TCCCATTCATCCCTGATGGAATGTTATAGTTAGTTAATCATTTGTATTTTCTAGT 1797  
 QY 1971 TGTATGTGATGTCTGTAAATAGTATATATTTGGCTTACAAATCCGTAAAGATG 2030  
 DB 1798 TGTATGTGATGTCTGTAAATAGTATATATTTGGCTTACAAATCCGTAAAGATG 1857  
 QY 2031 TTTGTCAATTTGAAATCTTAATGCCAAGTAAATGCAATGCTTTGGAAATTTGAAAGAT 2090  
 DB 1858 TTTGTCAATTTGAAATCTTAATGCCAAGTAAATGCAATGCTTTGGAAATTTGAAAGAT 1917  
 QY 2091 GGTTTATCTTTGAGAGCAAAATATGTTCATTAATATGCTTTGATTTGTCATATCAAG 2150  
 DB 1918 GGTTTATCTTTGAGAGCAAAATATGTTCATTAATATGCTTTGATTTGTCATATCAAG 1977  
 QY 2151 AATATGATGAACGTCTCAAAACCTGTTTACGCTACTCTGTAAAGAGGAGCCGGTTTGG 2210  
 DB 1978 AATATGATGAACGTCTCAAAACCTGTTTACGCTACTCTGTAAAGAGGAGCCGGTTTGG 2037  
 QY 2211 GAGGACATTCGATCGCTGCGCAAGTGTCTGTAAAGTGTCTTTAACTGAGAGGCG 2270  
 DB 2038 GAGGACATTCGATCGCTGCGCAAGTGTCTGTAAAGTGTCTTTAACTGAGAGGCG 2097  
 QY 2271 TAACTCAAAATACCTTTTAACTGATCTCTATATTAATATGAGGACAGTATGCTCTTA 2330  
 DB 2098 TAACTCAAAATACCTTTTAACTGATCTCTATATTAATATGAGGACAGTATGCTCTTA 2157  
 QY 2331 C 2331  
 DB 2158 C 2158

RESULT 10  
 HSU28833 2173 bp mRNA linear PRI 19-APR-2000  
 LOCUS Homo sapiens Down syndrome candidate region protein (DSCR1) mRNA,  
 DEFINITION complete cds.  
 ACCESSION U28833  
 VERSION U28833.2 GI:7596916  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2173)  
 Puentes,J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and  
 Estivill,X.  
 TITLE A new human gene from the Down syndrome critical region encodes a  
 proline-rich protein highly expressed in fetal brain and heart  
 JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)  
 MEDLINE 96121593  
 PUBMED 8595418  
 REFERENCES 2 (bases 1 to 2173)  
 Estivill,X.  
 AUTHORS Direct Submission

JOURNAL Submitted (09-JUN-1995) Xavier Estivill, Molecular Genetics  
 Department, Cancer Research Institute, Hospital Duran i Reynals,  
 Avia, Castelldefels Km 2.7, Hospitalet, Barcelona, Catalonia 08907,  
 Spain  
 3 (bases 1 to 2173)  
 REFERENCE  
 AUTHORS Estivill,X.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-APR-2000) Xavier Estivill, Molecular Genetics  
 Department, Cancer Research Institute, Hospital Duran i Reynals,  
 Avia, Castelldefels Km 2.7, Hospitalet, Barcelona, Catalonia 08907,  
 Spain  
 REMARK Sequence update by submitter  
 COMMENT On Apr 19, 2000 this sequence version replaced gi:1125051.  
 FEATURES  
 source  
 location/Qualifiers  
 1..2173  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.1-q22.2"  
 1..2173  
 /gene="DSCR1"  
 49..564  
 /gene="DSCR1"  
 /codon\_start=1  
 /product="down syndrome candidate region protein 1"  
 /protein\_id="AAB81557.1"  
 /db\_xref="GI:1125052"  
 /translation="MYYAKFESLFRDYDITFOYFKSKRYRINSPNSAADRLQ  
 LKITEFGKEMKLYFAQLIHGSHLAPRPDQFLISPPAPVGMKQVEDATPVIN  
 YDLVIAISKLGPEKYEELHAAITDTPSVVVHVCSDQEEBEMERMRPKRIQT  
 RRPETPIHLS"  
 2158  
 /gene="DSCR1"  
 polyA\_site  
 BASE COUNT 606 a 431 c 503 g 633 t  
 ORIGIN  
 Query Match 90.1%; Score 2101; DB 9; Length 2173;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 231 GCCAATTTGAGTCCCTCTTGAAGCGTATGACAGACATCACTTCACTATTTTAAAG 290  
 DB 58 GCCAATTTGAGTCCCTCTTGAAGCGTATGACAGACATCACTTCACTATTTTAAAG 117  
 QY 291 AGCTTCAACGAGTGAATTAATCTTCAGCAACCCCTCTCCGACAGAGTCCAGGCTC 350  
 DB 118 AGCTTCAACGAGTGAATTAATCTTCAGCAACCCCTCTCCGACAGAGTCCAGGCTC 177  
 QY 351 CAGCTCATAGACTGAGTTTCTGGGAAAGAAATGAATGTTATTTTGTCAAGCTTA 410  
 DB 178 CAGCTCATAGACTGAGTTTCTGGGAAAGAAATGAATGTTATTTTGTCAAGCTTA 237  
 QY 411 CACATGGAAGTCAACCTGCTCCGCCAATCCAGCAAGCAAGTTCTGATCTCCCT 470  
 DB 238 CACATGGAAGTCAACCTGCTCCGCCAATCCAGCAAGCAAGTTCTGATCTCCCT 297  
 QY 471 CCGGCTCTCCGCACTGGGATGGAACAAGTGAAGTGCAGCCCACTAAACTAT 530  
 DB 298 CCGGCTCTCCGCACTGGGATGGAACAAGTGAAGTGCAGCCCACTAAACTAT 357  
 QY 531 GATCTCTTATATCCATCTTCCAAGCTGGGCGCAGGGAAGATGATGATTCACCGCAGCG 590  
 DB 358 GATCTCTTATATCCATCTTCCAAGCTGGGCGCAGGGAAGATGATGATTCACCGCAGCG 417  
 QY 591 ACTGACACCACTCCCAAGCTGTGTCTCATGTATGAGAGTGAAGATGAAGAGAGGAA 650  
 DB 418 ACTGACACCACTCCCAAGCTGTGTCTCATGTATGAGAGTGAAGATGAAGAGAGGAA 477  
 QY 651 GAAGAGGAATGGAAGATGAAGAGACCTTAAGCCAAAATTAATCCAGACAGAGAGCGG 710  
 DB 478 GAAGAGGAATGGAAGATGAAGAGACCTTAAGCCAAAATTAATCCAGACAGAGAGCGG 537  
 QY 711 GAGTACAGCCGATCACTCACTGATGAGCAAGGAGAGAGAGATTCGAATCAT 770

```

Db 538 GAGTACACGCCGATCACCTCAGCTGAACTGGCACGCGACGAGGACGCAATTCRAATCAT 597
QY 771 ACTCAGCGGAGGAATCTTTTACTGGAGGTGGCTGGTCAACGACTTCTTCGGAGGTGCA 830
Db 598 ACTCAGCGGAGGAATCTTTTACTGGAGGTGGCTGGTCAACGACTTCTTCGGAGGTGCA 657
QY 831 GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTGTCAGAGAGAAATCAAGCCCGTG 890
Db 658 GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTGTCAGAGAGAAATCAAGCCCGTG 717
QY 891 TCCCTCTGTTCTAATGCTGCACACAGTACTGTTTCATGGCACCGCGGAATGACTTGGGC 950
Db 718 TCCCTCTGTTCTAATGCTGCACACAGTACTGTTTCATGGCACCGCGGAATGACTTGGGC 777
QY 951 CAATCACTGAGTTTGTGTGATCGCACAGGACATTTGGGACTGTCTTGAGAAACAGAT 1010
Db 778 CAATCACTGAGTTTGTGTGATCGCACAGGACATTTGGGACTGTCTTGAGAAACAGAT 837
QY 1011 AATGATAGTGTGTTTGTACTTCTTCTCTGTTAGTGGTCTGTCTGTCGCAAGGCGAGTT 1070
Db 838 AATGATAGTGTGTTTGTACTTCTTCTCTGTTAGTGGTCTGTCTGTCGCAAGGCGAGTT 897
QY 1071 GATCAGTGAGCTCAGGAGAGACTTCTGTTTCTAAGTGGCTCTCAGGCGCACTCTCTA 1130
Db 898 GATCAGTGAGCTCAGGAGAGACTTCTGTTTCTAAGTGGCTCTCAGGCGCACTCTCTA 957
QY 1131 CTGGTAGGAGAGTACACAGGAAGCGCTAGTCAGAGAGAGTGTGAAACAGCAGC 1190
Db 958 CTGGTAGGAGAGTACACAGGAAGCGCTAGTCAGAGAGAGTGTGAAACAGCAGC 1017
QY 1191 AATGCAATGTGGAATTTGTAGCGTTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGCATG 1250
Db 1018 AATGCAATGTGGAATTTGTAGCGTTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGCATG 1077
QY 1251 TATATTACTGATTACAGACTAACTTTGTCGTATATAAAGTTACACCGTTGTGTTT 1310
Db 1078 TATATTACTGATTACAGACTAACTTTGTCGTATATAAAGTTACACCGTTGTGTTT 1137
QY 1311 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAGTATACCTTTCCAGATTCTCG 1370
Db 1138 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAGTATACCTTTCCAGATTCTCG 1197
QY 1371 GATTCTCGACTCTTTGCAACAGCAGCTGCTTGGCGAACTCTTCTCGAATGCAATTCACTC 1430
Db 1198 GATTCTCGACTCTTTGCAACAGCAGCTGCTTGGCGAACTCTTCTCGAATGCAATTCACTC 1257
QY 1431 AGCATCCCAACCGTGCACAGTGTAACTTGTGCTTTTGCAGAAAGAGTTGATCTGAAAT 1490
Db 1258 AGCATCCCAACCGTGCACAGTGTAACTTGTGCTTTTGCAGAAAGAGTTGATCTGAAAT 1317
QY 1491 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGATAGCAGTTTCACTGCGCAACTTTTAGT 1550
Db 1318 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGATAGCAGTTTCACTGCGCAACTTTTAGT 1377
QY 1551 GGGTGAGAAATTTAGTTAGTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTT 1610
Db 1378 GGGTGAGAAATTTAGTTAGTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTT 1437
QY 1611 TGTGTGTGTTTCTATCATGATCATATAGCAAAATCTTTTGGAAACTGTTGGTGTGAGA 1670
Db 1438 TGTGTGTGTTTCTATCATGATGAATCATAGCAAAATCTTTTGGAAACTGTTGGTGTGAGA 1497
QY 1671 TAGTTGGTCTTTTACCACGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1730
Db 1498 TAGTTGGTCTTTTACCACGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1557
QY 1731 TATATTCACTCTGTTGTACATCTGGTGAAAGATATGSCAGTGGCAGACTAAGATGTAT 1790
Db 1558 TATATTCACTCTGTTGTACATCTGGTGAAAGATATGSCAGTGGCAGACTAAGATGTAT 1617
QY 1791 TAACTTACCTGTGAAATCATATCTGTTAGGAAAGCTGTTCCCATGCTTAACAGGACTTGA 1850

```

```

Db 1618 TAACCTACCTGTGAATCATATGTTGTAGGAAAAAGCTGTTCCCATGCTTAACAGGACTTGA 1677
QY 1851 ATTCAAAGCATGTCAAAGTGGATAGTAGTCTGTGGCGATATGAGAGGATGCAAGTGCCTT 1910
Db 1678 ATTCAAAGCATGTCAAAGTGGATAGTAGTCTGTGGCGATATGAGAGGATGCAAGTGCCTT 1737
QY 1911 TCCCCATTCATCTCTCATGTAAGTGTATATGTTTACATGCTTAACTTTGTAATTTTTTCTAGT 1970
Db 1738 TCCCCATTCATCTCTCATGTAAGTGTATATGTTTACATGCTTAACTTTGTAATTTTTTCTAGT 1797
QY 1971 TGTAAATGTGTATGCTGTGTAATAGGTATATATATTTTGGCCTTACAAATACCGTAACAATG 2030
Db 1798 TGTAAATGTGTATGCTGTGTAATAGGTATATATTTTGGCCTTACAAATACCGTAACAATG 1857
QY 2031 TTGTGTCATTTTGAATACATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2090
Db 1858 TTGTGTCATTTTGAATACATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1917
QY 2091 GGTGTTTATTTCTTTGAGAAGCAATATGTTTGCATTAATGCTTTTGAATTTTGAAGAT 2150
Db 1918 GGTGTTTATTTCTTTGAGAAGCAATATGTTTGCATTAATGCTTTTGAATTTTGAAGAT 1977
QY 2151 AAATGATTGAACGTTCTCAAAACCTGTTTACGTTACTTGGTAAGAGGAGCGGTTTGG 2210
Db 1978 AAATGATTGAACGTTCTCAAAACCTGTTTACGTTACTTGGTAAGAGGAGCGGTTTGG 2037
QY 2211 GAGAGACCATTCGATCGCTGCTCAAGTGTCTTCTTAAAGTGTCTTTAACTTGAAGAGC 2270
Db 2038 GAGAGACCATTCGATCGCTGCTCAAGTGTCTTCTTAAAGTGTCTTTAACTTGAAGAGC 2097
QY 2271 TAACTTCAAAATACATTTTTTTTAACTGCAATCTATATAAATGAGGACAGATGCTCCTTA 2330
Db 2098 TAACTTCAAAATACATTTTTTTTAACTGCAATCTATATAAATGAGGACAGATGCTCCTTA 2157
QY 2331 C 2331
Db 2158 C 2158

```

AK055845 2679 bp mRNA linear PRI 01-AUG-2002  
 Homo sapiens CDNA FLJ31283 fis, clone KIDNE2006820, highly similar  
 to DOWN SYNDROME CRITICAL REGION PROTEIN 1.  
 AK055845  
 AK055845 1 GI:16550675  
 Oligo capping; fis (full insert sequence).  
 Homo sapiens kidney cDNA to mRNA, clone\_lib:KIDNE2  
 Clone:KIDNE2006820.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,  
 Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,  
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,  
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
 WagaTsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
 Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2679)  
 Isogai,T., Otsuki,T. and Sugiyama,T.  
 Direct Submission  
 Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;  
 Research Association for Biotechnology (RAB); cDNA library  
 construction; Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; Clone selection for full insert sequencing; RAB and HRI.

# FEATURES

source

Location/Qualifiers -

1. 2679  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="KIDNE200820"  
/clone="KIDNE2"  
/clone="KIDNE2"  
/note="Cloning vector: pME18SFlu3"

BASE COUNT 739 a 540 c 631 g 769 t

ORIGIN

Query Match 89.8%; Score 2093.6; DB 9; Length 2679;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2096; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

230 GGGCAATTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCCTTCAGTATTTAA 289  
580 GGGCAATTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCCTTCAGTATTTAA 639  
290 GAGCTTCAACGAGTCAGATAAATTTCAGCAACCCCTTCTCCGACAGATGCCAGGCT 349  
640 GAGCTTCAACGAGTCAGATAAATTTCAGCAACCCCTTCTCCGACAGATGCCAGGCT 699  
350 CGAGTGCATAGACTGATTTCTGGGAAAGGAAATGAAATTTATTTGGTCCAGACCTT 409  
700 CGAGTGCATAGACTGATTTCTGGGAAAGGAAATGAAATTTATTTGGTCCAGACCTT 759  
410 ACACATGGAAGTCACACCTGGCTCCGCAAAATCCAGACAGATTTCTGATCTCCCC 469  
760 ACACATGGAAGTCACACCTGGCTCCGCAAAATCCAGACAGATTTCTGATCTCCCC 819  
470 TCCGCTCTCCGCAAGTGGATGGAAACAAGTGGAAAGTCCGACCTCTTAACCTA 529  
820 TCCGCTCTCCGCAAGTGGATGGAAACAAGTGGAAAGTCCGACCTCTTAACCTA 879  
530 TGAATCTTATATGCTATTCCTCCAAAGTGGGGCCAGGGGAAAGTATGACACGAGC 589  
880 TGAATCTTATATGCTATTCCTCCAAAGTGGGGCCAGGGGAAAGTATGACACGAGC 939  
590 GACTGACACCACTCCCAAGTGTGTGTCATGTATGTGAGATGATMAAGAAAGAGA 649  
940 GACTGACACCACTCCCAAGTGTGTGTCATGTATGTGAGATGATMAAGAAAGAGA 999  
650 AGAAGAGAAATGGAAGAAATGAGAGAACTTAAGCCAAATTTATCCAGACCGAGAGGCC 709  
1000 AGAAGAGAAATGGAAGAAATGAGAGAACTTAAGCCAAATTTATCCAGACCGAGAGGCC 1059  
710 GGAATACACCGCGCATCCACTCAGTGAATGAGACCGAGAGAGAGCATTCCTCAATCA 769  
1060 GGAATACACCGCGCATCCACTCAGTGAATGAGACCGAGAGAGAGCATTCCTCAATCA 1119  
770 TACTCAGCGGAGGAATCTTTTAACTGTGAGGTGCTGTCAGACTTCTTGGAGGTGC 829  
1120 TACTCAGCGGAGGAATCTTTTAACTGTGAGGTGCTGTCAGACTTCTTGGAGGTGC 1179  
830 AGCCGAGATCGGGGTGGAGAAATCCCAAGTTCATGTCTCAAGAGAAATCAAGGCCGT 889  
1180 AGCCGAGATCGGGGTGGAGAAATCCCAAGTTCATGTCTCAAGAGAAATCAAGGCCGT 1239  
890 GTCCCTGTTCTTAATGCTGACACCAAGTACTGTTTCAATGACACCGGGAAATGAATTGGG 949  
1240 GTCCCTGTTCTTAATGCTGACACCAAGTACTGTTTCAATGACACCGGGAAATGAATTGGG 1299  
950 CCAATCACTGACTTTGTGTGATGCAACAGGACATTTGGGACTGTCTTGAAGAAACGA 1009  
1300 CCAATCACTGACTTTGTGTGATGCAACAGGACATTTGGGACTGTCTTGAAGAAACGA 1359  
1010 TAAATAGATGTTTGTGATCTTCTTTCTGTGAGTCTGTGCTGCAAGAGGACAGT 1069

1360 TAAATAGATGTTTGTGATCTTCTTTCTGTGAGTCTGTGCTGCAAGAGGACAGT 1419  
1070 TGAATGAGTGCACAGAGAGACTTCTGTGTTCTAAGTGGCTGCAAGGACCACTCTG 1129  
1420 TGAATGAGTGCACAGAGAGACTTCTGTGTTCTAAGTGGCTGCAAGGACCACTCTG 1479  
1130 ACTGATGAGAGAGTACCAAGAGAGCCGCTTATGTCAGAGAGAGTGTGAAACAGCAG 1189  
1480 ACTGATGAGAGAGTACCAAGAGAGCCGCTTATGTCAGAGAGAGTGTGAAACAGCAG 1539  
1190 CAATGCAATGGAAGAAATGAGAGGTTCTTCTTCCCATGTTCTCAATGTTGTGCAAT 1249  
1540 CAATGCAATGGAAGAAATGAGAGGTTCTTCTTCCCATGTTCTCAATGTTGTGCAAT 1599  
1250 GTATATTTACTGATTTACAGACCTAACCTTTGTTGATATTAAGTTACACCGTTGTT 1309  
1600 GTATATTTACTGATTTACAGACCTAACCTTTGTTGATATTAAGTTACACCGTTGTT 1659  
1310 TTACATCTTTTGGGAGACCGCAAGAGGTTTGGAAACGATACCTTTCCAGATTTCTC 1719  
1660 TTACATCTTTTGGGAGACCGCAAGAGGTTTGGAAACGATACCTTTCCAGATTTCTC 1719  
1370 GGAATTCGACTTTTGGCAACAGCACTTGGCGAACTCTTCTGGAATGCAATTCAT 1429  
1720 GGAATTCGACTTTTGGCAACAGCACTTGGCGAACTCTTCTGGAATGCAATTCAT 1779  
1430 CAGATCCCAACCGTGCACGTTAACTTGTGCTTTTGGCAAGAAAGTTGATCTGAAT 1489  
1780 CAGATCCCAACCGTGCACGTTAACTTGTGCTTTTGGCAAGAAAGTTGATCTGAAT 1839  
1490 TCCCTGAGAAATTTAGGTTTATACATTCAGAGAAATGAGTTTCACTGCAACTTTTAA 1549  
1840 TCCCTGAGAAATTTAGGTTTATACATTCAGAGAAATGAGTTTCACTGCAACTTTTAA 1899  
1550 TGGGTGAGAAATTTAGTTTAAAGTTTGGATGGACCTTCACTTCTGTTCTTTT 1609  
1900 TGGGTGAGAAATTTAGTTTAAAGTTTGGATGGACCTTCACTTCTGTTCTTTT 1959  
1610 ATGTGTGTTTCTATATACATGATATGACCCAAATCTTTTGGAAACGTTGGTGGAG 1669  
1960 ATGTGTGTTTCTATATACATGATATGACCCAAATCTTTTGGAAACGTTGGTGGAG 2019  
1670 ATGTGTGTTTCTTAAACCCAGAAAGATGAGATGAGATGAGATGAGATGAGATGAG 1729  
2020 ATGTGTGTTTCTTAAACCCAGAAAGATGAGATGAGATGAGATGAGATGAGATGAG 2079  
1730 ATATATTCATACCTGTTGTAACCTTGGGTGAAAGATGAGATGAGATGAGATGAGATG 2139  
2080 ATATATTCATACCTGTTGTAACCTTGGGTGAAAGATGAGATGAGATGAGATGAGATG 2189  
1790 TTAACCTACCTGTAATCATGTTGTGAGAAACGTTCCATGTCATACAGAGACTTG 1849  
2140 TTAACCTACCTGTAATCATGTTGTGAGAAACGTTCCATGTCATACAGAGACTTG 2199  
1850 AATTCAAGCATGTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1909  
2200 AATTCAAGCATGTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 2259  
1910 TTCCCATTCATCTCTGATGAGAAATGTTTATCTAGATTAACTTTTCTGAG 1969  
2260 TTCCCATTCATCTCTGATGAGAAATGTTTATCTAGATTAACTTTTCTGAG 2319  
1970 TTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2029  
2320 TTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2379  
2030 GTTGTGATTTTGAATTTCTTAAATGCAATGCAATGCAATGCAATGCAATGCAATG 2089  
2380 GTTGTGATTTTGAATTTCTTAAATGCAATGCAATGCAATGCAATGCAATGCAATG 2439  
2090 TGGTTTATCTTTGAGAGCAATATGTTGATTAATGCTTTGATTTGATTTGATTTG 2449  
2440 TGGTTTATCTTTGAGAGCAATATGTTGATTAATGCTTTGATTTGATTTGATTTG 2499

QY 2150 GAAATTGATTGAACGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTG 2209  
Db 2500 GAAATTGATTGAACGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTG 2559  
QY 2210 GGAGAGACCATTTGCATCGCTGCTCCAAAGTGTCTTCTTTAAAGTCTTTTAAATCTGGAGAGG 2269  
Db 2560 GGAGAGACCATTTGCATCGCTGCTCCAAAGTGTCTTCTTTAAAGTCTTTTAAATCTGGAGAGG 2619  
QY 2270 CTAACCTCAAAATACATTTTAACTGCAATCTATAATAAATGGCAGATGCTCCTT 2329  
Db 2620 CTAACCTCAAAATATTTTAACTGCAATCTATAATAAATGGCAGATGCTCCTT 2679

RESULT 12  
AR034241  
LOCUS AR034241 2174 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869318.  
ACCESSION AR034241  
VERSION AR034241.1 GI:5949846  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 2174)  
Palleja,X.Estivill., Fuentes,J.Jose. and Pritchard,M.  
AUTHORS Gene sequence of the Down syndrome critical region of human  
TITLE chromosome 21, identified by a new 'Alu-splicing PCR' technique,  
coding for a proline-rich protein (DSCR1) highly expressed in  
foetal brain and in heart and method for characterizing it  
JOURNAL Patent: US 5869318-A 1 09-FEB-1999;  
FEATURES Location/Qualifiers  
Source 1..2174  
BASE COUNT 605 a 436 c 502 g 628 t 3 others  
ORIGIN

Query Match 88.3%; Score 2059.2; DB 6; Length 2174;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2081; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 231 GCCAATTTGAGTCCCTTTAGAGCTATGCAAGGACATCACCTTTTCAGTATTTTAAG 290  
Db 58 GCCAATTTGAGTCCCTTTAGAGCTATGCAAGGACATCACCTTTTCAGTATTTTAAG 117  
QY 291 AGCTTCAAACGAGTCAGATAAATCTTCAACACCCCTTCTCCGAGCAGATGCCAGGCTC 350  
Db 118 AGCTTCAAACGAGTCAGATAAATCTTCAACACCCCTTCTCCGAGCAGATGCCAGGCTC 177  
QY 351 CAGCTGATAAAGCTGAGTTCTTGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 410  
Db 178 CAGCTGATAAAGCTGAGTTCTTGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 237  
QY 411 CACATAGGAAGCTCACACCTGGCTCCGCCAATCCAGACAGAGTTCTGATCTCCCT 470  
Db 238 CACATAGGAAGCTCACACCTGGCTCCGCCAATCCAGACAGAGTTCTGATCTCCCT 297  
QY 471 CCGCCTCTCCGCACTGGGATGGAAACAAGTGGAGATGGCAACCCAGTCAATAACTAT 530  
Db 298 CCGCCTCTCCGCACTGGGATGGAAACAAGTGGAGATGGCAACCCAGTCAATAACTAT 357  
QY 531 GATCTCTTATATGCCATCTCAAGCTGGGCGAGGGGAAAGTATGAATTGCAAGGCG 590  
Db 358 GATCTCTTATATGCCATCTCAAGCTGGGCGAGGGGAAAGTATGAATTGCAAGGCG 417  
QY 591 ACTGACACCACTCCAGAGCTGTGTGCTCAATGATGAGAGTATCAAGAGAGGAA 650  
Db 418 ACTGACACCACTCCAGAGCTGTGTGCTCAATGATGAGAGTATCAAGAGAGGAA 477  
QY 651 GAAGAGGAATGGAAAGATGAGGAGACTTAAGCCAAAATTTATCCAGACGAGGCG 710  
Db 478 GAAGAGGAATGGAAAGATGAGGAGACTTAAGCCAAAATTTATCCAGACGAGGCG 537

QY 711 GAGTACACGCGCGATCCACCTCAGCTGAACCTGGCA CGCGACGAGGAGCGCATTCCAAATCAT 770  
Db 538 GAGTACACGCGCGATCCACCTCAGCTGAACCTGGCA CGCGACGAGGAGCGCATTCCAAATCAT 597  
QY 771 ACTCACGGGAGGAATCTTTTACTGTGGAGGTGGTGTGTCAGACCTTCTTCGAGGTGGCA 830  
Db 598 ACTCACGGGAGGAATCTTTTACTGTGGAGGTGGTGTGTCAGACCTTCTTCGAGGTGGCA 657  
QY 831 GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAAGAGAAATCAAGGCCG 890  
Db 658 GCCGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAAGAGAAATCAAGGCCG 717  
QY 891 TCCCTCTGTTCTTAATGCTGCGACACCACTTACTGTTCAATGGCACCCGCGGAATCACTTGG 950  
Db 718 TCCCTCTGTTCTTAATGCTGCGACACCACTTACTGTTCAATGGCACCCGCGGAATCACTTGG 777  
QY 951 CAATCACTGAGTTTGTGGTGTGTCGACCAAGGACATTTTGGGAGTGTCTTGAGAAACAGAT 1010  
Db 778 CAATCACTGAGTTTGTGGTGTGTCGACCAAGGACATTTTGGGAGTGTCTTGAGAAACAGAT 837  
QY 1011 AATGATAGTGTGTTGTACTTGTCTTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 1070  
Db 838 AATGATAGTGTGTTGTACTTGTCTTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 897  
QY 1071 GATCAGTGAAGTTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGCGAGGGCCACTTCTTA 1130  
Db 898 GATCAGTGAAGTTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGCGAGGGCCACTTCTTA 957  
QY 1131 CTGGTAGGAAGAGTTCACAGGAGCGCCCTAGTGCAGAGAGGTTGTGAAACAGCAGC 1190  
Db 958 CTGGTAGGAAGAGTTCACAGGAGCGCCCTAGTGCAGAGAGGTTGTGAAACAGCAGC 1017  
QY 1191 AATGCAATGTCGAAATGTTAGCTGTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGATG 1250  
Db 1018 AATGCAATGTCGAAATGTTAGCTGTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGATG 1077  
QY 1251 TATATTACTGATTTTACAAGACTAACCTTTGTTTGTATATAAAGTTACACCGTTGTTT 1310  
Db 1078 TATATTACTGATTTTACAAGACTAACCTTTGTTTGTATATAAAGTTACACCGTTGTTT 1137  
QY 1311 TACATCTTTTGGGAAGCGCAAGAGCTTTGGAAACGATATCACCTTTCCAGATTTCTCG 1370  
Db 1138 TACATCTTTTGGGAAGCGCAAGAGCTTTGGAAACGATATCACCTTTCCAGATTTCTCG 1197  
QY 1371 GATTTCTCGACTCTTTGCAACAGCAGCTTGTGGGAGCTCTTCTGGAATGCAATCACTC 1430  
Db 1198 GATTTCTCGACTCTTTGCAACAGCAGCTTGTGGGAGCTCTTCTGGAATGCAATCACTC 1257  
QY 1431 AGCATCCCAACCGTGCACGCTGTAACCTTGTGCTTTTGGAAAGAGTTGATCTGAAAT 1490  
Db 1258 AGCATCCCAACCGTGCACGCTGTAACCTTGTGCTTTTGGAAAGAGTTGATCTGAAAT 1317  
QY 1491 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATAGCAGTTTCACTGCAACTTTAGT 1550  
Db 1318 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATAGCAGTTTCACTGCAACTTTAGT 1377  
QY 1551 GGGTGGAAATTTTGTAGTTTGGATCGGAGCTTGTGGGAGCTTCTGTTGTTCTTTA 1610  
Db 1378 GGGTGGAAATTTTGTAGTTTGGATCGGAGCTTGTGGGAGCTTCTGTTGTTCTTTA 1437  
QY 1611 TGTGGTGGTTTCTATACATGAATCATAGCCAAACCTTTTGGAAACCTTGTGTTGAGA 1670  
Db 1438 TGTGGTGGTTTCTATACATGAATCATAGCCAAACCTTTTGGAAACCTTGTGTTGAGA 1497  
QY 1671 TAGTGTGTTCTTTTACCCACGAGAGACATCAAGATACACTTGTAAATAAGCTGATGACA 1730  
Db 1498 TAGTGTGTTCTTTTACCCACGAGAGACATCAAGATACACTTGTAAATAAGCTGATGACA 1557  
QY 1731 TATATTCTATCTGTTGTACACTTGGGTGAAAGATGAGCAGTGGAGACTAAGATGAT 1790  
Db 1558 TATATTCTATCTGTTGTACACTTGGGTGAAAGATGAGCAGTGGAGACTAAGATGAT 1617  
QY 1791 TAACTTACCTGTGTAATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAAACAGGACTTGA 1850

```

Db 1618 TAACCTACCTGGAATCATATGTTAGAGAAAAGCTGTCCCAATCTCAACAGACATTGA 1677
Qy 1851 ATTCAAGCATGTCAGATGATAGATCTGTGCGATATGAGGAGTGCAGTCCCTT 1910
Db 1678 ATTCAGAGCATGTCAGATGATAGATCTGTGCGATATGAGGAGTGCAGTCCCTT 1737
Qy 1911 TCCCATTCATTCCTGATGAGAAATTTGTTACTAGCTTAACATTTGTTTCTAGT 1970
Db 1738 TCCCATTCATTCCTGATGAGAAATTTGTTACTAGCTTAACATTTGTTTCTAGT 1737
Qy 1971 TGTATGTCATGTCGTGTAATAGGATATATTTGGCTTACAAATACCTTAACAATG 2030
Db 1798 TGTATGTCATGTCGTGTAATAGGATATATTTGGCTTACAAATACCTTAACAATG 1857
Qy 2031 TTTGTCATTTGGAATATCTTAATGCCAGTAAACATGATGCTTTGGAAATTTGGAAT 2090
Db 1858 TTTGTCATTTGGAATATCTTAATGCCAGTAAACATGATGCTTTGGAAATTTGGAAT 1917
Qy 2091 GGTATATTCCTTTGAGAGCAATATGTTGCTTAATAGCTTTGTCATATCAAG 2150
Db 1918 GGTATATTCCTTTGAGAGCAATATGTTGCTTAATAGCTTTGTCATATCAAG 1977
Qy 2151 AATGATGTAAGAGCTTCTCAACCTGTTACGGTACTTGTGAAGAGGAGCGGTTGG 2210
Db 1978 AATGATGTAAGAGCTTCTCAACCTGTTACGGTACTTGTGAAGAGGAGCGGTTGG 2037
Qy 2211 GAGAGCATTCGATGCTT-GTCCAGTGTCTTGTATAGCTTTTAACTCGAGAGG 2269
Db 2038 GAGAGCATTCGATGCTT-GTCCAGTGTCTTGTATAGCTTTTAACTCGAGAGG 2097
Qy 2270 CTAACTCCAAAATATCTTTTAACTGATCTATATAATAGGAGCAGTATGCTCTT 2329
Db 2098 CTAACTCCAAAATATCTTTTAACTGATCTATATAATAGGAGCAGTATGCTCTT 2157
Qy 2330 AC 2331
Db 2158 AC 2159

```

```

RESULT 13
AY007090/c 1792 bp mRNA linear PRI 31-AUG-2000
LOCUS Homo sapiens clone TCCCIA00112 mRNA sequence.
DEFINITION AY007090
ACCESSION AY007090.1 GI:9955979
VERSION AY007090.1 GI:9955979
KEYWORDS FLI-CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1792)
AUTHORS Anderson, B., Wentland, M.A., Ricefrence, J.Y., Liu, W., and Gibbs, R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474
REFERENCE 2 (bases 1 to 1792)
AUTHORS Yu, W., Anderson, B., Morley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
Ricefrence, J.Y., Wentland, M.A., Lennon, G., and Gibbs, R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
PUBMED 9110174
REFERENCE 3 (bases 1 to 1792)
AUTHORS Zhou, J., Yu, W., Tang, H., Mei, G., Tsang, Y.T.M., Bouck, J., Gibbs, R.A.,
and Margolin, J.F.
TITLE Direct Submision
JOURNAL Submitted (24-JUN-2000) Human Genome Sequencing Center and Texas
Children's Cancer Center, Baylor College of Medicine, Houston, TX
77030, USA
COMMENT The clone request should be directed to Dr. J. Margolin at
Pediatric-Hematology & Oncology, Texas Children's Center

```

```

FEATURES
source
102514, Houston, Texas 77030, USA. Telephone: 713-770-4583
margolin@bcm.tmc.edu.
Location/Qualifiers
1..1792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCCCIA00112"
/clone_lib="constructed by Y.T.M. Tsang"
/dev_stage="infant"
/note="from patient with acute lymphoblastic leukemia"
BASE COUNT 547 a 428 c 335 g 482 t
ORIGIN
Query Match 76.8%; Score 1790.4; DB 9; Length 1792;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1791; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 536 CTTATATGCATCTCCAAAGCTGGGGCCAGGAAAGTATGATATTCACGACGACTGA 595
Db 1792 CTTATATGCATCTCCAAAGCTGGGGCCAGGAAAGTATGATATTCACGACGACTGA 1733
Qy 596 CACCACCTCCAGCCTGTGTGTCATGTATGTAGATGATCAAGAGAGAGAGAGAG 655
Db 1732 CACCACCTCCAGCCTGTGTGTCATGTATGTAGATGATCAAGAGAGAGAGAGAG 1673
Qy 656 GGAATGGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715
Db 1672 GGAATGGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1613
Qy 716 CAGCGCATTCACCTCAGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
Db 1612 CAGCGCATTCACCTCAGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1553
Qy 776 CGGAGAGAGATCTTTACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 835
Db 1552 CGGAGAGAGATCTTTACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1493
Qy 836 GATCGGGGGGAGAGAAATCCCAAGTTCATGTTCTCAGAGAGAGAGAGAGAG 895
Db 1492 GATCGGGGGGAGAGAAATCCCAAGTTCATGTTCTCAGAGAGAGAGAGAGAG 1433
Qy 896 TTGTTCTATATGTCGACACACAGTTACTGTTATGAGCACCAGGAAATGAG 955
Db 1432 TTGTTCTATATGTCGACACACAGTTACTGTTATGAGCACCAGGAAATGAG 1373
Qy 956 ACTGAGTTGTGTGATGAGCACAAGAGATTTGGGACTGCTTGAGAAACAGATATGA 1015
Db 1372 ACTGAGTTGTGTGATGAGCACAAGAGATTTGGGACTGCTTGAGAAACAGATATGA 1313
Qy 1016 TAGTGTGTTGCTGCTGTTCTTTCTGTAGAGTTCTGTCTGTGCCAAGGGCAGTTGATCA 1075
Db 1312 TAGTGTGTTGCTGCTGTTCTTTCTGTAGAGTTCTGTCTGTGCCAAGGGCAGTTGATCA 1253
Qy 1076 GTGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135
Db 1252 GTGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1193
Qy 1136 AGGAAGAGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195
Db 1192 AGGAAGAGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1133
Qy 1196 AATGTGAATTTAGAGCTTTCTTTCTTCCATGTTCTCATGTTGTGATGATAT 1255
Db 1132 AATGTGAATTTAGAGCTTTCTTTCTTCCATGTTCTCATGTTGTGATGATAT 1073
Qy 1256 TACTGATTTCAAGAGATTAACCTTTGTTGCTGATTAAGTAAACCGTTGTTTACAT 1315
Db 1072 TACTGATTTCAAGAGATTAACCTTTGTTGCTGATTAAGTAAACCGTTGTTTACAT 1013
Qy 1316 CTTTGGAGAGCAGAGAAACGTTTGGAAACGATACCTTTCCAGATTCGAGATTC 1375
Db 1012 CTTTGGAGAGCAGAGAAACGTTTGGAAACGATACCTTTCCAGATTCGAGATTC 953

```



Qy	1376	TCGACTCTTTGCAACAGCACATTGCTTGGCGAACTCTTCTCGGAATGCATTCCTCAGCAT	1435
Db	952	TCGACTCTTTGCAACAGCACATTGCTTGGCGAACTCTTCTCGGAATGCATTCCTCAGCAT	893
Qy	1436	CCCAACCGTCGACGTTAACTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATTCCTCT	1495
Db	892	CCCAACCGTCGACGTTAACTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATTCCTCT	833
Qy	1496	GTAGAAATTTAGCTTTATACAAATTCAGAGAAATAGCAGTTTTCACCTGCCAACTTTTGTGGGT	1555
Db	832	GTAGAAATTTAGCTTTATACAAATTCAGAGAAATAGCAGTTTTCACCTGCCAACTTTTGTGGGT	773
Qy	1556	AGAAATTTTAGTTTGTGGATCGGACCTCAGTTTCTGTGTGTTCTTTTATGTGG	1615
Db	772	AGAAATTTTAGTTTGTGGATCGGACCTCAGTTTCTGTGTGTTCTTTTATGTGG	713
Qy	1616	TGTTTCTTATACATCAATCATAGCCAAAACCTTTTGGAAAATGTTTGGTTGAGATAGTT	1675
Db	712	TGTTTCTTATACATCAATCATAGCCAAAACCTTTTGGAAAATGTTTGGTTGAGATAGTT	653
Qy	1676	GGTTCTTTTACCCGCAAGACATCAAGATACACTTGTGAATAAAGCTGATAGCATATAT	1735
Db	652	GGTTCTTTTACCCGCAAGACATCAAGATACACTTGTGAATAAAGCTGATAGCATATAT	593
Qy	1736	TCATACCTGTTGTACACTTCGGGTGAAAGCTATGGCAGTGGGAGCTAAGATGATTAAAC	1795
Db	592	TCATACCTGTTGTACACTTCGGGTGAAAGCTATGGCAGTGGGAGCTAAGATGATTAAAC	533
Qy	1796	TACCTGTGAATCATATGTTGTAGAAAAGCTGTTCCCATGCTTAACAGGACTTGAATTC	1855
Db	532	TACCTGTGAATCATATGTTGTAGAAAAGCTGTTCCCATGCTTAACAGGACTTGAATTC	473
Qy	1856	AAGCATGTCAGTGGATAGTAGATCTGTGGCATATGAGAGGATGACGTGCTTTTCCCG	1915
Db	472	AAGCATGTCAGTGGATAGTAGATCTGTGGCATATGAGAGGATGACGTGCTTTTCCCG	413
Qy	1916	ATTCATTCCTGATGGAATGTTTATCTAGGTTAAACATTTGTAATTTTTTCTAGTTGTAA	1975
Db	412	ATTCATTCCTGATGGAATGTTTATCTAGGTTAAACATTTGTAATTTTTTCTAGTTGTAA	353
Qy	1976	TGCTGATGTCGGTAATAGGTATTATATTTTGGCTTACAATACCGTAACAAATGTTCT	2035
Db	352	TGCTGATGTCGGTAATAGGTATTATATTTTGGCTTACAATACCGTAACAAATGTTCT	293
Qy	2036	CATTTTGAATACTTAATGCCAAGTAAACAATGCATGCTTTGGAAATTTGGAAGATGGTT	2095
Db	292	CATTTTGAATACTTAATGCCAAGTAAACAATGCATGCTTTGGAAATTTGGAAGATGGTT	233
Qy	2096	TATCTTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTGTTTCATATCAAGAAAT	2155
Db	232	TATCTTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTGTTTCATATCAAGAAAT	173
Qy	2156	GATTCAAAGTCTTCAAAACCTGTTTACCGTACTTGGTAAGGGAGCCGGTTGGAGAG	2215
Db	172	GATTCAAAGTCTTCAAAACCTGTTTACCGTACTTGGTAAGGGAGCCGGTTGGAGAG	113
Qy	2216	ACCATTCATCCGCTGTCCAAGTGTCTTGTGTAAGTGCTTTTAACTGCGAGGCTAAC	2275
Db	112	ACCATTCATCCGCTGTCCAAGTGTCTTGTGTAAGTGCTTTTAACTGCGAGGCTAAC	53
Qy	2276	TCAAAATACTTTTTTAACTGCATCTATATAAATGGCCACAGTATGCTCC	2327
Db	52	TCAAAATACTTTTTTAACTGCATCTATATAAATGGCCACAGTATGCTCC	1
RESULT 14	AP000326/c		
LOCUS	AP000326	83762 bp	DNA linear PRI 20-NOV-1999
DEFINITION	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone: P0125H6, complete sequence.		
ACCESSION	AP000326		
VERSION	AP000326.1	GI:4835695	
KEYWORDS	HTG.		

QY 1103 CTAAAGGCTGTCAGAGGCACTCTCTACTGTAGAAAGAGTACACAGAGAGCCGCT 1162  
 DB 9707 CTAAAGGCTGTCAGAGGCACTCTCTACTGTAGAAAGAGTACACAGAGAGCCGCT 9648  
 QY 1163 AGTGCAGAGAGTGTGAAAACAGCAGCAGTGCATGTGAAAATTGTGACGCTTCTTC 1222  
 DB 9647 AGTGCAGAGAGTGTGAAAACAGCAGCAGTGCATGTGAAAATTGTGACGCTTCTTC 9588  
 QY 1223 TTCCCTCATGTTCTCATGTTGTGACATGATATATCTGATTTACAAAGCTAACCTTTGT 1282  
 DB 9587 TTCCCTCATGTTCTCATGTTGTGACATGATATATCTGATTTACAAAGCTAACCTTTGT 9528  
 QY 1283 CGTATATAAGTTACACCGTGTGTTTGTATATCTTTTGGAAAGCCAGAAAGCGTTGG 1342  
 DB 9527 CGTATATAAGTTACACCGTGTGTTTGTATATCTTTTGGAAAGCCAGAAAGCGTTGG 9468  
 QY 1343 AAAAGCTATACCTTTCCAGATTCCTCGGATTCCTGACCTTTGGAAAGCAGACTTGTG 1402  
 DB 9467 AAAAGCTATACCTTTCCAGATTCCTCGGATTCCTGACCTTTGGAAAGCAGACTTGTG 9408  
 QY 1403 CGGAACCTTCTCGAATGATTCATCTAGCATCCCAACCGTGCAGACGTATACCTTGTG 1462  
 DB 9407 CGGAACCTTCTCGAATGATTCATCTAGCATCCCAACCGTGCAGACGTATACCTTGTG 9348  
 QY 1463 CTTTTCGAAAAGAGTGTGATCTGAAATTCCTCTGTAGAAATTAGCTTATACATTCAGAG 1522  
 DB 9347 CTTTTCGAAAAGAGTGTGATCTGAAATTCCTCTGTAGAAATTAGCTTATACATTCAGAG 9288  
 QY 1523 AATGACAGTTTCACTGCGCACTTTAGTGGTGAATAATTTAGTTAGTGTGGAT 1582  
 DB 9287 AATGACAGTTTCACTGCGCACTTTAGTGGTGAATAATTTAGTTAGTGTGGAT 9228  
 QY 1583 CGGACCTCAGTTCTGTGTTCTTTTATATGTGTGTTCTATACATGATTCATAGCCAA 1642  
 DB 9227 CGGACCTCAGTTCTGTGTTCTTTTATATGTGTGTTCTATACATGATTCATAGCCAA 9168  
 QY 1643 AAACCTTTTGGAAAGCTGTGTTGAGATAGTTGTTCTTTTACCAGCAAGACATCAA 1702  
 DB 9167 AAACCTTTTGGAAAGCTGTGTTGAGATAGTTGTTCTTTTACCAGCAAGACATCAA 9108  
 QY 1703 GATACACTGTAAATTAAGTGTATGATCATATATTCATCTGTGTATCACTTGGTGA 1762  
 DB 9107 GATACACTGTAAATTAAGTGTATGATCATATATTCATCTGTGTATCACTTGGTGA 9048  
 QY 1763 AGTATGCAAGTGTGAGACTAAGATATTAACCTACCTGTGATCATATGTGTGAGAA 1822  
 DB 9047 AGTATGCAAGTGTGAGACTAAGATATTAACCTACCTGTGATCATATGTGTGAGAA 8988  
 QY 1823 AGCTGTTCCATGTCTTAACAGGACTTGAATCAAGCATGTCAAGTGTGATGATCTG 1882  
 DB 8987 AGCTGTTCCATGTCTTAACAGGACTTGAATCAAGCATGTCAAGTGTGATGATCTG 8928  
 QY 1887 TGGGCAATGAGAGGAGTACAGTGCCTTTCCCATTCATCTGTGATGAAATTGTATCT 1942  
 DB 8927 TGGGCAATGAGAGGAGTACAGTGCCTTTCCCATTCATCTGTGATGAAATTGTATCT 8868  
 QY 1943 AGCTTAACATTTGTAATTTTCTAGTTGTAATGTATGTCGTGTAATAGTATAT 2002  
 DB 8867 AGCTTAACATTTGTAATTTTCTAGTTGTAATGTATGTCGTGTAATAGTATAT 8808  
 QY 2003 ATTTTGGCCTTCAATACGCTAACAATGTGTGATTTGAAATATCTTAATGCCAAGTAA 2062  
 DB 8807 ATTTTGGCCTTCAATACGCTAACAATGTGTGATTTGAAATATCTTAATGCCAAGTAA 8748  
 QY 2063 CAATGATGCTTTGAAATTGGAAGATGTTTATCTTTGGAAGCAAAATATGTTGC 2122  
 DB 8747 CAATGATGCTTTGAAATTGGAAGATGTTTATCTTTGGAAGCAAAATATGTTGC 8688  
 QY 2123 ATTAATATGCTTATATGTTCTATATCAAGAAATGATGGAAGCTTCAAAACCTGTTTAC 2182  
 DB 8687 ATTAATATGCTTATATGTTCTATATCAAGAAATGATGGAAGCTTCAAAACCTGTTTAC 8628

QY 2183 GGTACTTGTGTAAGAGGAGCCGTTTGGAGAGACCATTCATCCCTGTCCAAAGTCTTC 2242  
 DB 8627 GGTACTTGTGTAAGAGGAGCCGTTTGGAGAGACCATTCATCCCTGTCCAAAGTCTTC 8568  
 QY 2243 TTGTTAAGCTCTTTTAATCTGAGAGGCTTAACCTCAAAATPACCTTTTAACTGATTC 2302  
 DB 8567 TTGTTAAGCTCTTTTAATCTGAGAGGCTTAACCTCAAAATPACCTTTTAACTGATTC 8508  
 QY 2303 ATATATAAGGAGCAGTATGCTCTTAC 2331  
 DB 8507 ATATATAAGGAGCAGTATGCTCTTAC 8479  
 RESULT 15  
 AP000054/c 100000 bp DNA linear PRI 20-NOV-1999  
 LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, segment 25/28,  
 DEFINITION  
 ACCESSION AP000054  
 VERSION AP000054.1 GI:3132364  
 KEYWORDS HTG.  
 SOURCE Homo sapiens DNA, clone:245P17-f4A4f\_4.  
 ORGANISM Homo sapiens  
 BUKARYOTA; METAZOA; CHORDATA; CRANIATA; VERTEBRATA; EUTELEOSTOMI;  
 MAMMALIA; EUTHERIA; PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 REFERENCE  
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.  
 TITLE Homo sapiens genomic DNA, chromosome 21q  
 JOURNAL Published only in Database (1998)  
 REFERENCE  
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAY-1998) Masahira Hattori, Kitasato University,  
 Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,  
 Sagamihara 228, Japan (E-mail: hattori@igc.ims.u-tokyo.ac.jp,  
 Tel:0427-78-9732, Fax:0427-78-9561)  
 COMMENT This sequence is conducted by Kitasato University JST sequencing  
 Laboratory as a JST sequencing team.  
 Principal Investigator: Yoshiyuki Sakaki Ph.D.  
 Phone: +81-3-5449-5622, Fax: +81-3-5449-5445,  
 sakaki@igc.ims.u-tokyo.ac.jp  
 Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The  
 sequence is submitted by Human Genome Sequencing in ALIS project of  
 JST  
 Japan Science and Technology Corporation (JST)  
 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our sequence  
 archive Web site (<http://www-alis.tokyo.jst.go.jp/HGS/top.html>)  
 or send email to [webmaster@www-alis.tokyo.jst.go.jp](mailto:webmaster@www-alis.tokyo.jst.go.jp).  
 FEATURES  
 source location/Qualifiers  
 1..100000  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /map="21q22.1"  
 /clone="245P17-f4A4f\_4"  
 BASE COUNT 28599 a 21687 c 22052 g 27662 t  
 ORIGIN  
 Query Match 75.8%; Score 1767.4; DB 9; Length 100000;  
 Best local Similarity 99.9%; Pred. No. 0;  
 Matches 1769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 563 AGGGAAAAGTGAATTGACGACGAGATGTGACACACTCCAGGCGTGTGCTCATGT 622  
 DB 27996 AGGGAAAAGTGAATTGACGACGAGATGTGACACACTCCAGGCGTGTGCTCATGT 27937  
 QY 623 ATGAGAGTGATCAAG 662  
 DB 27936 ATGAGAGTGATCAAG 27877  
 QY 683 GCCAAAATTATCCAGACGAGAGAGCGGAGATACAGCCCATCACTCAAGCTGAAGCTGG 742



Db	27876	GCCAAATAATATCCAGACAGGAGCGGAGTACACGCCGATCCACCTCAGCTGAAGTGG	27817
QY	743	CACGCGACGAGGACGCAATCCCAATCATACTCAGCGGAGGAATCTTTTACTGTGAGGTG	802
Db	27816	CACGCGACGAGGACGGCAATTCCAAATCATACTCAGCGGAGGAATCTTTTACTGTGAGGTG	27757
QY	803	GCTGGTCAAGCAATCTTCTCGAGGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAAGTCA	862
Db	27756	GCTGGTCAAGCAATCTTCTCGAGGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAAGTCA	27697
QY	863	TGTTGCTCAGAGAGAAATCAAGCCCGTGTCCCTTGTCTTAATGTGCACACCACTTACT	922
Db	27696	TGTTGCTCAGAGAGAAATCAAGCCCGTGTCCCTTGTCTTAATGTGCACACCACTTACT	27637
QY	923	GTTTCATGGCACCCGGGAATGACTCTGGGCCAAATCACTGAGTTTGTGGTGTATCGCA	982
Db	27636	GTTTCATGGCACCCGGGAATGACTCTGGGCCAAATCACTGAGTTTGTGGTGTATCGCA	27577
QY	983	CATTTGGGACTGTCTTGAGAAACAGATAATGATAGTGTTTTGTAATCTGTCTTTCTTGG	1042
Db	27576	CATTTGGGACTGTCTTGAGAAACAGATAATGATAGTGTTTTGTAATCTGTCTTTCTTGG	27517
QY	1043	TAGGTTCTGTCTGTGCCAAGGCGAGTTGATCAGTGAGCTCAGAGAGAGCTTCTGTGTT	1102
Db	27516	TAGGTTCTGTCTGTGCCAAGGCGAGTTGATCAGTGAGCTCAGAGAGAGCTTCTGTGTT	27457
QY	1103	CTAAGTGGCTGCAGGGCCACTCTCTACTGCTAGGAAGAGGTACACAGAGAACGCCCT	1162
Db	27456	CTAAGTGGCTGCAGGGCCACTCTCTACTGCTAGGAAGAGGTACACAGAGAACGCCCT	27397
QY	1163	AGTGACAGAGAGGTTGTGAAAAACAGCAGCAATGCAAATGTGGAATTTGTAGCGTTTCTTTTC	1222
Db	27396	AGTGACAGAGAGGTTGTGAAAAACAGCAGCAATGCAAATGTGGAATTTGTAGCGTTTCTTTTC	27337
QY	1223	TTCCCTCATGTTCTCATGTTTGTGTCATGTATTTACTGATTTTACAAGCACTAACCTTGTT	1282
Db	27336	TTCCCTCATGTTCTCATGTTTGTGTCATGTATTTACTGATTTTACAAGCACTAACCTTGTT	27277
QY	1283	CGTATATAAGGTTACACGGTTGTTGTTTACATCTTTTGGGAAGCCAGGAAGCGTTGG	1342
Db	27276	CGTATATAAGGTTACACGGTTGTTGTTTACATCTTTTGGGAAGCCAGGAAGCGTTGG	27217
QY	1343	AAAAAGTATACCTTTTCCAGATTTCCGATTTCTCGATCTTGTGCAACAGCACTTGCTTG	1402
Db	27216	AAAAAGTATACCTTTTCCAGATTTCTCGATTTCTCGATCTTGTGCAACAGCACTTGCTTG	27157
QY	1403	CGGAATCTTCTCTGGAATGCAATTCACCTCAGCATCCCCCAACCGTGAACGTTGAACCTTG	27097
Db	27156	CGGAATCTTCTCTGGAATGCAATTCACCTCAGCATCCCCCAACCGTGAACGTTGAACCTTG	27037
QY	1463	CTTTTGCAAAAGAGTTGATCTGMAATTCCTCTGTAGAATTTAGCTTATACAAATCAGAG	1522
Db	27096	CTTTTGCAAAAGAGTTGATCTGMAATTCCTCTGTAGAATTTAGCTTATACAAATCAGAG	27037
QY	1523	AATAGCAGTTTTCATGCGCAACTTTTATGTTGGGTGAGAAATTTTATGTTTGGGAT	1582
Db	27036	AATAGCAGTTTTCATGCGCAACTTTTATGTTGGGTGAGAAATTTTATGTTTGGGAT	26977
QY	1583	CGGACCTCAGTTTCTGTGTTCTTTTATGTGGTGGTTTCTATACATGAATCATAGCCAA	1642
Db	26976	CGGACCTCAGTTTCTGTGTTCTTTTATGTGGTGGTTTCTATACATGAATCATAGCCAA	26917
QY	1643	AAACTTTTTTGGAAACTGTGTGGTTGAGATAGTTGGTTCTTTTACCCCAAGAGACATCAA	1702
Db	26916	AAACTTTTTTGGAAACTGTGTGGTTGAGATAGTTGGTTCTTTTACCCCAAGAGACATCAA	26857
QY	1703	GATACACTGTGTAATAAAGCTGATGCATATATTCATACCTGTGTACACTTTGGGTGAAA	1762
Db	26856	GATACACTGTGTAATAAAGCTGATGCATATATTCATACCTGTGTACACTTTGGGTGAAA	26797
QY	1763	AGTATGGCAGTGGGACCTAAGATGTTTAACTCCTGTGAAATCATATGTTGTAGGAAA	1822

Db	26796	AGTATGGCAGTGGGAGCTAAGATGTATTAACTACCTGTGATCATATGTTGTAGGAAA	26737
Qy	1823	AGCTGTTCCCATGTCTTACAGGACTTGAATTCAAAGCATGTCAAAGTATAGTAGATCTG	1882
Db	26736	AGCTGTTCCCATGTCTTAAAGGACTTGAATTCAAAGCATGTCAAAGTATAGTAGATCTG	26677
Qy	1883	TGGCGATATGAGAGGATGCAAGTGCTTTCCCAATTCATCTCGATGGAAATGTTATACT	1942
Db	26676	TGGCGATATGAGAGGATGCAAGTGCTTTCCCAATTCATCTCGATGGAAATGTTATACT	26617
Qy	1943	AGGTTAAATTTGTAATTTTTTCTAGTGTGAATGTATGTCTGCTAAATAGGTAATTAT	2002
Db	26616	AGGTTAAATTTGTAATTTTTTCTAGTGTGAATGTATGTCTGCTAAATAGGTAATTAT	26557
Qy	2003	ATTTTGGCCCTTACAAATCCGTAACAAATGTTTGTCTATTTGAAATCTTAAATGCCAAGTAA	2062
Db	26556	ATTTTGGCCCTTACAAATCCGTAACAAATGTTTGTCTATTTGAAATCTTAAATGCCAAGTAA	26497
Qy	2063	CAATGCATGCTTTTGGAAATTTGGAAGATGGTGTATTCTTTTGAGAAGCAAAATATGTTTGC	2122
Db	26496	CAATGCATGCTTTTGGAAATTTGGAAGATGGTGTATTCTTTTGAGAAGCAAAATATGTTTGC	26437
Qy	2123	ATTTAAATGCTTTGATTTGTTTCATATCAAGAAATGATTTGAAGCTTCTCAAAACCTGTTTAC	2182
Db	26436	ATTTAAATGCTTTGATTTGTTTCATATCAAGAAATGATTTGAAGCTTCTCAAAACCTGTTTAC	26377
Qy	2183	GGTACTTGGTAAGAGGGAGCGGTTTGGGAGAGCAATTCGATCGCTGTCCAAAGTGTTC	2242
Db	26376	GGTACTTGGTAAGAGGGAGCGGTTTGGGAGAGCAATTCGATCGCTGTCCAAAGTGTTC	26317
Qy	2243	TTGTTAAAGTGCTTTTAAACTGGAGAGGCTAACCTTCAAAATACTTTTTTAACTGCATCTCT	2302
Db	26316	TTGTTAAAGTGCTTTTAAACTGGAGAGGCTAACCTTCAAAATATTTTTTTAACTGCATCTCT	26257
Qy	2303	ATAATAATGGGCACATATGCTCCTTAC	2331
Db	26256	ATAATAATGGGCACATATGCTCCTTAC	26228

Search completed: December 14, 2002, 17:10:26  
Job time : 4511.35 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:06:10 ; Search time 662.336 Seconds  
(without alignments)  
14597.874 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atgagatttaggacttag.....caccgatccaccttagctga 597

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_estc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pin:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_Other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	587.4	98.4	1158	14	BQ895506
2	566.6	94.9	753	13	BI148584
3	532.6	89.2	710	10	BB617325
4	532.2	89.1	665	13	BI219142
5	532	89.1	950	12	BF782633
6	530	88.8	980	12	BQ296537

7	517	86.6	711	13	BG915568
8	516	86.4	827	13	BI102432
9	506.2	84.8	1507	11	AK010696
10	504.4	84.5	559	9	AA200984
11	472.2	79.1	837	9	AU124628
12	470.6	78.8	939	9	AL546617
13	470.6	78.8	1041	13	BM450020
14	469.6	78.7	718	9	AU131040
15	469	78.6	931	9	AL543576
16	467.4	78.3	946	14	BQ278576
17	460.6	77.2	740	13	BI463566
18	459.6	77.0	885	9	AL544755
19	456.2	76.4	520	14	BQ749142
20	455.8	76.3	906	9	AL544313
21	455.4	76.3	939	9	AL554686
22	450.2	75.4	662	13	BI464521
23	445.4	74.6	1078	13	BM541636
24	441.8	74.0	769	13	BI767955
25	434.4	72.8	501	12	BF385073
26	433.8	72.7	710	12	BG475986
27	432.6	72.5	811	12	BG570239
28	431.4	72.3	751	9	AL576189
29	408.2	68.4	890	9	AL538796
30	407.8	68.3	828	9	AL536447
31	406.2	68.0	956	9	AL556803
32	404.2	67.7	705	9	AL550372
33	403	67.5	1047	12	BG287042
34	401.6	67.3	1014	9	AL559594
35	401.4	67.2	627	13	BI859506
36	399.2	66.9	875	12	BG574693
37	397.8	66.6	937	12	BE795722
38	395.6	66.3	867	14	BQ427531
39	395	66.2	923	9	AL551657
40	382	64.0	986	12	BG828534
41	375.8	62.9	859	12	BF203629
42	375.6	62.9	535	14	BM832622
43	374	62.6	822	12	BE895008
44	372.2	62.3	525	9	AL037245
45	369.4	61.9	727	13	BI327875

## ALIGNMENTS

RESULT 1  
BQ895506  
LOCUS BQ895506 1158 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT\_8762889 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6329174  
S, mRNA sequence.  
ACCESSION BQ895506  
VERSION BQ895506.1 GI:22287520  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1158)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.  
cDNA Library Preparation: ResGen, Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM13781 row: i column: 15  
High quality sequence start: 83  
High quality sequence stop: 577.

BG915568 602815815  
BI102432 602889619  
AK010696 Mus muscu  
AA200984 mul2e11.r  
AU124628 AU124628  
AL546617 AL546617  
BM450020 AGENCOURT  
AU131040 AU131040  
AL543576 AL543576  
BQ278576 AGENCOURT  
BI463566 603206210  
AL544755 AL544755  
BQ749142 UI-M-PD0-  
AL544313 AL544313  
AL54686 AL54686  
BI464521 603206117  
BM541636 AGENCOURT  
BI767955 603061320  
BF385073 602045671  
BG475986 602521111  
BG570239 602590641  
AL576189 AL576189  
AL538796 AL538796  
AL536447 AL536447  
AL556803 AL556803  
AL550372 AL550372  
BQ287042 602382338  
AL559594 AL559594  
BI859506 603385358  
BG574693 602596709  
BE795722 601590730  
BQ427531 AGENCOURT  
AL551657 AL551657  
BG828534 602752653  
BF203629 601865655  
BM832622 K-EST0107  
BE895008 601433720  
AL037245 DKFZP564P  
BI327875 602979461

## FEATURES

Location/Qualifiers  
1. 1158  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6329174"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccd; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by Resgen, Invitrogen Corp. Note: this is a NIH-MGC Library."

BASE COUNT 289 a 354 c 288 g 226 t 1 others  
ORIGIN

Query Match 98.4%; Score 587.4; DB 14; Length 1158;  
Best Local Similarity 99.0%; Pred. No. 6.3e-154;  
Matches 591; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ATGATTTTGGGACTTTAGTACATATTTAGCTCCCTGATTTGCTGTGGCAACGAT 60  
109 ATGCATTTTGGGACTTTAGTACATATTTAGCTCCCTGATTTGCTGTGGCAACGAT 168  
61 GATGCTTTCAGGCAAGTGTAGACAGGCGCAATTGAATCCCTTTCAGAACATATGAC 120  
169 GATGCTTTCAGGCAAGTGTAGACAGGCGCAATTGAATCCCTTTCAGAACATATGAC 228  
121 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTTGCCGATTAACCTTCAGCAAC 180  
229 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTTGCCGATTAACCTTCAGCAAC 288  
181 CCTTATCTGACCGCATGTCAGGCTGCGGCTGACAAAGCGAGTTCCTGGGGAAGAA 240  
289 CCTTATCTGACCGCATGTCAGGCTGCGGCTGACAAAGCGAGTTCCTGGGGAAGAA 348  
241 ATGAAGTTGATTTTGTCTGACACTTTCACATAGAAAGTTTCAACCTGCTGCCCAAT 300  
349 ATGAAGTTGATTTTGTCTGACACTTTCACATAGAAAGTTTCAACCTGCTGCCCAAT 408  
301 CCCGACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGTTGGCTGAAACAAAGTA 360  
409 CCCGACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGTTGGCTGAAACAAAGTA 468  
361 GAAGATGCCACCCCGCATTAATTAAGATTTTATATGCACTTCACAGCTGGGGCCA 420  
469 GAAGATGCCACCCCGCATTAATTAAGATTTTATATGCACTTCACAGCTGGGGCCA 528  
421 GGAGAGAGTATGAACTGATGACGAGCGACAGCCCACTCCCAATGTTGGTGTCCAGTG 480  
529 GGAGAGAGTATGAACTGATGACGAGCGACAGCCCACTCCCAATGTTGGTGTCCAGTG 588  
481 TGTGAGAGTATGAACTGATGACGAGCGAGAGAGAGAGAGATGAGAAAGAGAGCC 540  
589 TGTGAGAGTATGAACTGATGACGAGCGAGAGAGAGAGAGATGAGAAAGAGAGCC 648  
541 AAGCCCAAAATCATTCAGACGAGACCGGAGTACACACCGATTCACCTTAAGCTGA 597  
649 AAGCCCAAAATCATTCAGACGAGAGCGGAGTACACACCGATTCACCTTAAGCTGA 705

RESULT 2  
B1148584 753 bp mRNA linear EST 05-JUL-2001  
LOCUS B1148584 60291195F1 NCI\_GAP\_L19 Mus musculus cDNA IMAGE:5053246 5',  
DEFINITION mRNA Sequence.  
ACCESSION B1148584 GI:14608585  
VERSION B1148584.1  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 753)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strusberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNL11144 row: m column: 23  
High quality sequence Etop: 751.

FEATURES  
source Location/Qualifiers  
1. 753  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5053246"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI-GAP Library."

BASE COUNT 207 a 197 c 192 g 157 t  
ORIGIN

Query Match 94.9%; Score 566.6; DB 13; Length 753;  
Best Local Similarity 99.0%; Pred. No. 3.5e-148;  
Matches 591; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

1 ATGATTTTGGGACTTTAGTACATATTTAGCTCCCTGATTTGCTGTGGCAACGAT 60  
89 ATGCATTTTGGGACTTTAGTACATATTTAGCTCCCTGATTTGCTGTGGCAACGAT 148  
61 GATGCTTTCAGGCAAGTGTAGACAGGCGCAATTGAATCCCTTTCAGAACATATGAC 120  
149 GATGCTTTCAGGCAAGTGTAGACAGGCGCAATTGAATCCCTTTCAGAACATATGAC 207  
121 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTTGCCGATTAACCTTCAGCAAC 180  
208 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTTGCCGATTAACCTTCAGCAAC 267  
181 CCTTATCTGACCGCATGTCAGGCTGCGGCTGACAAAGCGAGTTCCTGGGGAAGAA 240  
268 CCTTATCTGACCGCATGTCAGGCTGCGGCTGACAAAGCGAGTTCCTGGGGAAGAA 327  
241 ATGAAGTTGATTTTGTCTGACACTTTCACATAGAAAGTTTCAACCTGCTGCCCAAT 300  
328 ATGAAG-TGATTTTGTCTGACACTTTCACATAGAAAGTTTCAACCTGCTGCCCAAT 386  
301 CCCGACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGTTGGCTGAAACAAAGTA 360  
387 CCCGACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGTTGGCTGAAACAAAGTA 446  
361 GAAGATGCCACCCCGCATTAATTAAGATTTTATATGCACTTCACAGCTGGGGCCA 420  
447 GAAGATGCCACCCCGCATTAATTAAGATTTTATATGCACTTCACAGCTGGGGCCA 506  
421 GGAGAGAGTATGAACTGATGACGAGCGACAGCCCACTCCCAATGTTGGTGTCCAGTG 480  
507 GGAGAGAGTATGAACTGATGACGAGCGACAGCCCACTCCCAATGTTGGTGTCCAGTG 566  
481 TGTGAGAGTATGAACTGATGACGAGCGAGAGAGAGAGATGAGAAAGAGAGCC 540  
567 TGTGAGAGTATGAACTGATGACGAGCGAGAGAGAGAGATGAGAAAGAGAGCC 626  
541 AAGCCCAAAATCATTCAGACGAGACCGGAGTACACACCGATTCACCTTAAGCTGA 597  
627 AAGCCCAAAATCATTCAGACGAGAGCGGAGTACACACCGATTCACCTTAAGCTGA 683



REFERENCE 1 (bases 1 to 665)  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLM11238 row: 0 column: 21  
 High quality sequence stop: 657.  
 Location/Qualifiers

FEATURES  
 source 1..665  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5099372"  
 /lab\_host="NCI CGAP L49"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."  
 Location/Qualifiers

BASE COUNT 192 a 170 c 158 g 145 t  
 ORIGIN

Query Match 89.1%; Score 532.2; DB 13; Length 665;  
 Best Local Similarity 97.3%; Pired. No. 1.5e-138; Indels 3; Gaps 3;  
 Matches 573; Conservative 0; Mismatches 13;

1 ATGATTTTGGGACTTACCTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 60  
 77 ATGATTTTGGGACTTACCTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 136  
 61 GATGCTTTCAGGAAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTACACATATGAC 120  
 137 GATGCTTTCAGGAAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTACACATATGAC 195  
 121 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTGTCGGATMAACTTCAGCAAC 180  
 196 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTGTCGGATMAACTTCAGCAAC 255  
 181 CCCTTATCTGAGCGGATGCGAGGCTGCGCTGCAACAAGCCGAGTTCCTGGGAAAGAA 240  
 256 CCCTTATCTGAGCGGATGCGAGGCTGCGCTGCAACAAGCCGAGTTCCTGGGAAAGAA 315  
 241 ATGAAGTTGATTTTGTCTCAACCTTACACATAGGAATTCACACCTGCTCCGCCAAT 300  
 316 ATGAAGTTGATTTTGTCTCAACCTTACACATAGGAATTCACACCTGCTCCGCCAAT 375  
 301 CCCGCAAAAGATTCTCATCTCCCTCCGGCTCTCTCCCTGGTGGTGAACCAAGTA 360  
 376 CCCGCAAAAGATTCTCATCTCCCTCCGGCTCTCTCCCTGGTGGTGAACCAAGTA 435  
 361 GAAGATGCCACCCCGCTCATTAATTAAGATCTTTTATATGCAATCCAGCTGGGGCCA 420  
 436 GACCAATTTATACCCCGCTCATTAATTAAGATCTTTTATATGCAATCCAGCTGGGGCCA 495  
 421 GAGAGAAATATGAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 496 GAGAGAAATATGAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555  
 481 TGTGAGATGAG 539  
 556 TGTGAGATGAG 615  
 540 CAAGCCCAAAATATATCAGACA-CGAGAGCCGAGTACACAGATCCA 587

Db 616 AAGCCCAAAATATATCAGACACCGAGAGCCGAGTACACAGATCCA 664

RESULT 5  
 LOCUS BF782633 950 bp mRNA linear EST 12-JAN-2001  
 DEFINITION 602107462P1 NCI CGAP Kid14 Mus musculus cDNA clone IMAGE:4236038  
 5', mRNA sequence.  
 ACCESSION BF782633  
 VERSION BF782633.1 GI:12087669  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 950)  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLM9843 row: k column: 15  
 High quality sequence stop: 701.  
 Location/Qualifiers

FEATURES  
 source 1..950  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4236038"  
 /lab\_host="NCI CGAP Kid14"  
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."  
 Location/Qualifiers

BASE COUNT 220 a 257 c 274 g 199 t  
 ORIGIN

Query Match 89.1%; Score 532; DB 12; Length 950;  
 Best Local Similarity 98.3%; Pired. No. 2.1e-138; Indels 5; Gaps 5;  
 Matches 590; Conservative 0; Mismatches 5;

1 ATGATTTTGGGACTTACCTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 60  
 76 ATGATTTTGGGACTTACCTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 134  
 61 GATGCTTTCAGGAAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTACACATATGAC 120  
 135 GATGCTTTCAGGAAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTACACATATGAC 194  
 121 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTGTCGGATMAACTTCAGCAAC 180  
 195 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTGTCGGATMAACTTCAGCAAC 254  
 181 CCCTTATCTGAGCGGATGCGAGGCTGCGCTGCAACAAGCCGAGTTCCTGGGAAAGAA 240  
 255 CCCTTATCTGAGCGGATGCGAGGCTGCGCTGCAACAAGCCGAGTTCCTGGGAAAGAA 314  
 241 ATGAAGTTGATTTTGTCTCAACCTTACACATAGGAATTCACACCTGCTCCGCCAAT 300  
 315 ATGAAGTTGATTTTGTCTCAACCTTACACATAGGAATTCACACCTGCTCCGCCAAT 373  
 301 CCCGCAAAAGATTCTCATCTCCCTCCGGCTCTCTCCCTGGTGGTGAACCAAGTA 360  
 374 CCCGCAAAAGATTCTCATCTCCCTCCGGCTCTCTCCCTGGTGGTGAACCAAGTA 433



Library constructed by Life Technologies. Investigators providing samples: Lohar Heminghaus/Priestella Furch, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 196 a 184 c 183 g 148 t

Query Match 86.6%; Score 517; DB 13; Length 711;  
Best Local Similarity 98.6%; Pred. No. 2.9e-134;  
Matches 553; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

22 TACAAATTTAGTCCCTGATGCTGTGGCAACGATGATGCTTTCAGCGAAAGTAG 81  
1 TACAAATTTAGTCCCTGATGCTGTGGCAACGATGATGCTTTCAGCGAAAGTAG 60  
82 ACCAGGCGCAATTTGAATCCCTCTTCAGAACATGACCAAGACCACTTCCAGTAT 141  
61 ACCA-GGCCAATTTGAATCCCTCTTCAGAACATGACCAAGACCACTTCCAGTAT 119  
142 TTTAAGAGCTTCAACGCTGCTCGGATAAATTGACCAACCCCTTATTCGAGCGATGCG 201  
120 TTTAAGAGCTTCAACGCTGCTCGGATAAATTGACCAACCCCTTATTCGAGCGATGCG 179  
202 AGGCTCGGCTGACAAAGCCGAGTCTCGGGAAGGAATGATGATTTTGTCTGAG 261  
180 AGGCTCGGCTGACAAAGCCGAGTCTCGGGAAGGAATGATGATTTTGTCTGAG 239  
262 ACTTTACATAGGAAGTTTACACCTGCTCGGCAATCCGCAACAAAGTTTCTGATC 321  
240 ACTTTACATAGGAAGTTTACACCTGCTCGGCAATCCGCAACAAAGTTTCTGATC 299  
322 TCCCTCTCGGCTCTCTCTCCGCTGGAAGGAAGGAAGTTTCTGATC 380  
300 TCCCTCTCGGCTCTCTCTCCGCTGGAAGGAAGGAAGTTTCTGATC 359  
381 AAATTCAGATCTTTTATATGCTCATCTTCCAGCTGGGCGCAGAGAGATGATGATGCA 440  
360 AAATTCAGATCTTTTATATGCTCATCTTCCAGCTGGGCGCAGAGAGATGATGATGCA 419  
441 TCGAGCGACAGACCCCTCCAGTGTGTGTGCTCAACCTGTGTGAGAGTACCAAGAA 500  
420 TCGAGCGACAGACCCCTCCAGTGTGTGTGCTCAACCTGTGTGAGAGTACCAAGAA 479  
501 TCGAGAGGAAGAGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 560  
480 TCGAGAGGAAGAGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 538  
561 ACGAGAGCCGAGATACACACC 581  
539 ACGAGAGCCGAGATACACACC 559

RESULT 8 827 bp mRNA linear EST 26-JUN-2001  
LOCUS B1102432 602889619P1 NCI CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5044612  
DEFINITION 5' mRNA sequence.

ACCESSION B1102432  
VERSION B1102432.1 GI:14553325  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 827)  
AUTHORS NIH-MGC hftp://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bcr-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM1122 row: f column: 05  
High quality sequence stop: 666.

FEATURES  
Source 1..827  
Location/Qualifiers

BASE COUNT 206 a 248 c 204 g 169 t

Query Match 86.4%; Score 516; DB 13; Length 827;  
Best Local Similarity 97.1%; Pred. No. 6.1e-134;  
Matches 579; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

1 ATGATTTTGGGACTTACATTTTACCTCCCTGATGCTGTGGCAACGAT 60  
76 ATGATTTTGGGACTTACATTTTACCTCCCTGATGCTGTGGCAACGAT 135  
61 GATGCTTACGCGAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGAT 120  
136 GATGCTTACGCGAAGTGAAGACCA-GGCCAATTTGAATCCCTCTTCAGAACATATGAT 194  
121 AAGGACCACTTTCAGATTTTAAAGCTTCAACGCTGCGGATTAACCTTACAGAAC 180  
195 AAGGACCACTTTCAGATTTTAAAGCTTCAACGCTGCGGATTAACCTTACAGAAC 254  
181 CCTTATCTGAGCGGAGTCCGAGCTGCGGCTGCAAGACCGAGTTCTGCGGAGAGAA 240  
255 CCTTATCTGAGCGGAGTCCGAGCTGCGGCTGCAAGACCGAGTTCTGCGGAGAGAA 314  
241 ATGAAGTTGATTTTGTCTCAGACTTACATAGGAAGTTTCAACCTGCTCGGCCAAT 300  
315 ATGAAGTTGATTTTGTCTCAGACTTACATAGGAAGTTTCAACCTGCTCGGCCAAT 374  
301 CCGGACAAACGTTCTGATCTCCCTCCGCTCTCTCTCCGCTGCTGCTGCTGCTGCTG 360  
375 CCGGACAAACGTTCTGATCTCCCTCCGCTCTCTCTCCGCTGCTGCTGCTGCTGCTG 434  
361 GAAGATGCGACCCCGCTCATTAATTTACGATCTTTATATGCTGCTCAAGCTGGGCGCA 420  
435 GAAGATGCGACCCCGCTCATTAATTTACGATCTTTATATGCTGCTCAAGCTGGGCGCA 494  
421 GGAGGAGATGATGAATGCTGATGCGAGACGACCCCACTCCCACTGCTGCTGCTGCTG 480  
495 GGAGGAGATGATGAATGCTGATGCGAGACGACCA-TCCAGTGTGTGTGTGTGTGTG 553  
481 TGTGAGAGTGAACCAAGAGATGAG--AGGAAGAGAGAGATG--GAGAGATGAAGAG 536  
554 TGTGAGAGTGAACCAAGAGATGAG--AGGAAGAGAGATGAG--GAGAGATGAAGAG 613  
537 ACCCAAGCCCAAAATCATCCAGACA-CGAGAGCCGAGATACACACCGATTCACCTT 591  
614 ACCCAAGCCCAAAATCATCCAGACA-CGAGAGCCGAGATACACACCGATTCACCTT 669

RESULT 9 1507 bp mRNA linear HTC 19-JAN-2002  
LOCUS AK010696 Mus musculus ES cells cDNA, RIKEN full-length enriched library,  
DEFINITION clone:2410048A02.Down syndrome critical region homolog 1 (human),  
full insert sequence.  
ACCESSION AK010696





[illegible]

BASE COUNT	164 a	150 c	149 g	136 t	
ORIGIN					
Query Match			84.5%	Score 504.4	DB 9; Length 599;
Best Local Similarity			99.6%	Pred. No. 9.3e-131;	
Matches 516; Conservative			0; Mismatches 1;	Indels 1;	Gaps 1;
Oy	4	GATTTTAGGACCTTACTACAAATTTAGTCCCTGATTTGCTGTGTGGCAACAGATGAT	63		
Db	83	GATTTTAGGACCTTACTACAAATTTAGTCCCTGATTTGCTGTGTGGCAACAGATGAT	142		
Oy	64	GTCTTCAGCGAAGTGAGACCCAGGCGCAATTTGGAATCCCTTTACAGACATATGACAA	123		
Db	143	GTCTTCAGCGAAGTGAGACCCAGGCGCAATTTGGAATCCCTTTACAGACATATGACAA	202		
Oy	124	GACACCACTTCCTCAGTATTTTAAAGCTTCAAAGTGTCCGGGATTAACCTTCAGAACCC	163		
Db	203	GACACCACTTCCTCAGTATTTTAAAGCTTCAAAGTGTCCGGGATTAACCTTCAGAACCC	262		
Oy	184	TTATCTGCAGCCGATGCGCAGGCTGGCTGCAGACAGACCGAGTTCTTGCGGAAAGAAATG	243		
Db	263	TTATCTGCAGCCGATGCGCAGGCTGGCTGCAGACAGACCGAGTTCTTGCGGAAAGAAATG	322		
Oy	244	AAAGTTGATTTTGTCTCAGACTTTTACATATGAGTTTCAACCTGGCTCCGCCAATCCC	303		
Db	323	AAAGTTGATTTTGTCTCAGACTTTTACATATGAGTTTCAACCTGGCTCCGCCAATCCC	381		
Oy	304	GACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGCTGGGAAACAGTAGAA	363		
Db	382	GACAAACAGTTCTCATCTCCCTCCGCTCTCTCCCTCCGCTGGGAAACAGTAGAA	441		
Oy	364	GATGCCACCCCGCTCATTAATTAGATCTTTTATATGCAATCTCCAGCTGGGCGCAGGA	423		
Db	442	GATGCCACCCCGCTCATTAATTAGATCTTTTATATGCAATCTCCAGCTGGGCGCAGGA	501		
Oy	424	GAGAGTATGAACTGCAATGCAAGCAAGCCCACTCCAGTGTGTGTCTCACTGTGT	483		
Db	502	GAGAGTATGAACTGCAATGCAAGCAAGCCCACTCCAGTGTGTGTCTCACTGTGT	561		
Oy	484	GAGAGTACCAAGCAAGTGAAGGAGGAGGAGGAGAT	521		
Db	562	GAGAGTACCAAGCAAGTGAAGGAGGAGGAGGAGAT	599		
RESULT 11					
AUI24628	AUI24628	837 bp	mRNA	linear	EST 01-AUG-2002
LOCUS	AUI24628	NT2RM4	Homo sapiens	cdna	clone NT2MM4000307 5', mRNA
DEFINITION	AUI24628				
ACCESSION	AUI24628				
VERSION	AUI24628.1	GI:10949344			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
	1 (bases 1 to 837)				
	Oca,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,				
	Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and				
	HRI human cDNA project (Oca,T., Wakamatsu,A., Ozawa,M., Ishii,S.,				
	Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki				
	Y., Sugano,S., Isegai,T.)				
	Unpublished (2000)				
	Contact: Takao Isegai				
	Genomics Laboratory				
	Helix Research Institute				
	1532-3 Yana, Kibaraizu, Chiba 293-0812, Japan				
	Tel: 81-438-52-3975				
	Fax: 81-438-52-3986				
	Email: genomcshri.co.jp				
	HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix				

```

*
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. .937
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000307"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM4"
/cell_line="NT2"
/note="Vector: pME18SPL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 241 a 201 c 207 g 186 t 2 others
ORIGIN
Query Match 79.1%; Score 472.2; DB 9; Length 837;
Best Local Similarity 88.1%; Pred. No. 1.2e-121;
Matches 526; Conservative 0; Mismatches 68; Indels 3; Gaps 1;
QY 1 ATGGATTTTAGGACCTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGGCAAAACGAT 60
Db 118 ATGCATTTTAGAACTTTAACTACAGTTTATAGCTCCCTGATTTGCTGTGGCAACAGT 177
QY 61 GATGCTTTCAGCGAAAGTCAGACACAGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120
Db 178 GATATCTTCAGCGAAAGTGAACACAGGGCCAAATTTGATCCCTCTTTAGGACGTATGAC 237
QY 121 AAGGACACACCTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAAACTTCAGCAAC 180
Db 238 AAGGACATCACCTTTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAAACTTCAGCAAC 297
QY 181 CCCTTATCTGCAGCGGATGCCAGGCTGCGGTGCACAGACGAGATTCCTGGGAAAGGAA 357
Db 298 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 420
QY 241 ATGAAGTTGTATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTCCGCCCAAT 300
Db 358 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTCCGCCCAAT 417
QY 301 CCCGACAACAGTTTCCTCATCTCCCTCCGCTCCGCTCTCTCCCTGGCTGGCAACAGTA 360
Db 418 CCAGCAACAGCTTTCTGATCTCCCTCCGCTCTCCGCTCCGCTGGGATGGAAACAGTG 477
QY 361 GAAGATGCCACCCCGTCATATAATTTAGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 420
Db 478 GAAGATGCCACCCCGTCATATAATTTAGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 537
QY 421 GGAGAGATGATGAATGCAATGCAGCAGACACCCACTCCAGTGTGGTGGTCCACGCTG 480
Db 538 GGGGAAAGATGATGAATGCAATGCAGCAGACACCCACTCCAGTGTGGTGGTCCATGTA 597
QY 481 TGTGAGATGACCAAGAGATGAGGAGAGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 598 TGTGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
QY 541 AAGCCAAAATTCATCCAGACACCGAGACCGGAGTACACACCGATCCACCTTACCTGA 597
Db 655 AAGCCAAAATTCATCCAGACACCGAGAGCGGAGTACACCGATCCACCTTACCTGA 711

RESULT 12
AL546617 939 bp mRNA linear EST 16-FEB-2001
LOCUS AL546617 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1029YJ21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL546617
VERSION AL546617
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 939)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. .939  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1029YJ21"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@liferetech.com URL :  
http://fulllength.invitrogen.com"

## FEATURES source

1. .939

263 a 219 c 238 g 218 t 1 others

Query Match 78.8%; Score 470.6; DB 9; Length 939;

Best Local Similarity 87.9%; Pred. No. 3.7e-121;

Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

QY 1 ATGGATTTTAGGACCTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGGCAAAACGAT 60

Db 53 ATGCATTTTAGAACTTTAACTACAGTTTATAGCTCCCTGATTTGCTGTGGCAACAGT 112

QY 61 GATGCTTTCAGCGAAAGTCAGACACAGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120

Db 113 GATATCTTCAGCGAAAGTGAACACAGGGCCAAATTTGATCCCTCTTTAGGACGTATGAC 172

QY 121 AAGGACACACCTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAAACTTCAGCAAC 180

Db 173 AAGGACATCACCTTTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAAACTTCAGCAAC 232

QY 181 CCCTTATCTGCAGCGGATGCCAGGCTGCGGTGCACAGACGAGATTCCTGGGAAAGGAA 240

Db 233 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 292

QY 241 ATGAAGTTGTATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTGGCAACAGTA 300

Db 293 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTGGCAACAGTA 352

QY 301 CCCGACAACAGTTTCCTCATCTCCCTCCGCTCCGCTCTCTCCCTGGCTGGCAACAGTA 360

Db 353 CCAGCAACAGCTTTCTGATCTCCCTCCGCTCTCCGCTCCGCTGGGATGGAAACAGTG 412

QY 361 GAAGATGCCACCCCGCTCATAAATTTAGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 420

Db 413 GAAGATGGACCCCACTCATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA 472

QY 421 GGAGAGATGATGAATGCAATGCAGCAGACACCCACTCCAGTGTGGTGGTCCACGCTG 480

Db 473 GGGGAAAGATGATGAATTTGCAACGAGGAGTGCACACCTCCAGCTGGTGGTCCATGTA 532

QY 481 TGTGAGATGACCAAGAGATGAGGAGAGAGAGAGAGAGAGAGAGAGATGAAGAGACCC 540

Db 533 TGTGAGATGATCAAG 589

QY 541 AAGCCAAAATTCATCCAGACACCGAGACCGGAGTACACACCGATCCACCTTACCTGA 597

Db 590 AAGCCAAAATTCATCCAGACACCGAGAGCGGAGTACACCGATCCACCTTACCTGA 646

RESULT	13
BMs50020	
LOCUS	BMs50020 1041 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGNCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
ACCESSION	BMs50020
VERSION	5, mRNA sequence.
KEYWORDS	BMs50020.1 GI:18499060
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryote, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/
COMMENT	Unpublished (1999)
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Email: cgapbs-remail.nih.gov
	Tissue Procurement: ATCC/DCTP
	cDNA Library Preparation: Life Technologies, Inc.
	DNA Sequencing by: Agencourt Bioscience Corporation
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LNL at:
	http://image.llnl.gov
	plate: LLM12204 row: d column: 10
	High quality sequence stop: 643.
FEATURES	Location/Qualifiers
source	1..1041
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5528025"
	/clone_1lb="NIH_MGC_72"
	/tissue_type="melanotic melanoma"
	/lab_host="DH10B (phage-resistant)"
	/notes="Organ: skin; Vector: pCMV-SORTC; Site 1: NotI;
	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
	Average insert size 2 kb. Library constructed by Life
	Technologies."
BASE COUNT	282 a 248 c 274 g 236 t 1 others
ORIGIN	
Query Match	78.8%; Score 470.6; DB 13; Length 1041;
Best Local Similarity	87.9%; Pred. No. 3,9e-121;
Matches 525; Conservative	0; Mismatches 69; Indels 3; Gaps 1
Oy	1 ATGATTTTAGGACTTAAGCTTAACAATTAAAGTCCTCGGTGGTGCGCAACGAT 60
Db	84 ATGATTTTAAACCTTTAAGCTTAAGCTTTAGCTTCCTGATGCTGTGTGGCAACG 143
Oy	61 GATGCTTCAGGGAAGGAGACGAGGCCAAATTTGATCCCTCTTCAGAACATATGAC 120
Db	144 GATATCTTCAGCGAAGAGAAACGAGGCCAAATTTGATCCCTCTTAGAGCGTATGAC 203
Oy	121 AAGGACACCACCTTCAGTATTTTAAGAGCTTCAAACCTGTCGGATTAACCTGAGCAC 180
Db	204 AAGGACATCACCTTCAGTATTTTAAGAGCTTCAAACGATGATTAACCTTCAGCAC 263
Oy	181 CCCTATCTCGAGCGGAGCCGAGGCTGGGCTGCACAGACCGAGTTCTGGGGAAGAA 240
Db	264 CCGTCTCCGAGCGAGATGCCAGGCTCCAGCTGCAATAAGACGATTTCTGGGAAGGAA 323
Oy	241 ATGAAGTGTATTTTGCTCAGACTTAACATAGAGAAGTTACACCTTGCTCCGCCAAT 300
Db	324 ATGAAGTATATTTTGCTCAGACTTAACATAGAGAAGCTCACCTGCTCCGCCAAT 383
Oy	301 CCCGACAAACAGTTCTCATCTCCCTCCGCGCTCCCTCCGTTGGCGGAAACAATA 360
Db	384 CCAGACAAACAGTTCTCATCTCCCTCCGCGCTCCCTCCGCGAGTGAAGAAACAATG 443
Oy	361 GAAGATGCCACCCCGCTCAATTAATTAGATCTTTTATGATGCATCTCCAAGCTGGAGCA 420
Db	444 GAAGATGCCACCCCGCTCAATTAATTAGATCTTTTATGATGCATCTCCAAGCTGGAGCA 503

Qy	421	GGAGAGAATGATTAACCTGCATGACGAGACAGACCCTCCAGTGTGTGGTCACAGTG	480
Ds	504	GCGGAAGAAAGTAAATTGACCGACGACTGACACCACTCCAGCTGTGTGCCATGTGA	563
Qy	481	TGTGAGAGTACCAAGAGATGAGCAGGAAGAGAGAGAGATGAGAGATGAAGAGACCC	540
Ds	564	TGTGAGAGTATCAAGAGAGAGAGAGAGAGAGAGAG--ATGAGAAAGATGAGAGACCT	620
Qy	541	AAGCCAAAATCATCCAGACACGAGACCGGAGTTACACACCGATCCACTTGTAGCTGA	597
Ds	621	AAGCCAAAATTTATCCAGACACGAGAGCGGAGATCACCGCATCCACTTCAGCTGA	677
<b>RESULT 14</b>			
LOCUS	AU131040	718 bp	mRNA linear EST 01-AUG-2002
DEFINITION	AU131040 NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA sequence.		
ACCESSION	AU131040		
VERSION	AU131040.1	GI:10991394	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T. HRI human cDNA project Unpublished (2000) Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..718 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="NT2RP3001895" /clone_id="NT2RP3" /cell_type="teratocarcinoma" /cell_line="NT2" /note="Vector: pME18SFJ3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"		
FEATURES	source		
BASE COUNT	217 a 174 c 169 g 155 t		
ORIGIN			
Query Match	78.7%; Score 469.6; DB 9; Length 718;		
Best Local Similarity	87.8%; Pred. No. 6,1e-121;		
Matches	524; Conservative 0; Mismatches 70; Indels 3; Gaps 1.		
Qy	1	ATGATATTTAGGAGCTTACTACTACAATTTTAGCTCCCTGATGCTGTGTGCMAACGAT	60
Ds	102	ATGATATTTTGAACCTTAACACAGTTTTAGCTCCCTGATGCTGTGTGCMAACGAT	161
Ds	61	GATGCTTTCAGCGAAGATGAGACCGAGGCCCAAATTTGAATCCTCTTCAGAACATATGAC	120
Ds	162	GATATCTTCAGCGAAGATGAGAACCGAGGCCCAAATTTGAATCCTCTTCAGAACATATGAC	221
Qy	121	AAGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATTAACCTTCAGCAAC	180
Ds	222	AAGACATCAACCTTTCAGTATTTTAAAGAGCTTCAAAAGTGTTCGAATTAACCTTCAGCAAC	281
Qy	181	CCCTTATGTGAGCGAGTCCAGGCTCGGCTGCACAGACCGAGTTCTTGAGGAAGAA	240



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.3524 Seconds  
(without alignments)  
513.170 Million cell updates/sec

Title: US-09-782-953-3  
 Perfect score: 1048  
 Sequence: 1 MEEVDLQDLSATIACHLP ..... RPKKIIQTRRPEYTFIHL 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

total number of hits satisfying chosen parameters: 262574

```

minimum DB seq length: 0
maximum DB seq length: 2000000000

```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```

Database :
1 : /cn2.6/pdata/1/iaa/5A COMB pep.*
2 : /cn2.6/pdata/1/iaa/5B COMB pep.*
3 : /cn2.6/pdata/1/iaa/6A COMB pep.*
4 : /cn2.6/pdata/1/iaa/6B COMB pep.*
5 : /cn2.6/pdata/1/iaa/PCTUS COMB pep.*
6 : /cn2.6/pdata/1/iaa/backfiles1 pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	835.5	79.7	171	2	US-08-665-040-2	Sequence 2, Appli
2	88	8.4	284	5	PCT-US94-09753-4	Sequence 4, Appli
3	84.5	8.1	719	4	US-09-641-741-28	Sequence 28, Appli
4	82	7.8	1128	4	US-09-641-741-30	Sequence 30, Appli
5	82	7.8	1128	4	US-09-660-482-8	Sequence 8, Appli
6	81	7.7	1128	1	US-08-111-339-2	Sequence 2, Appli
7	79	7.5	1132	1	US-08-446-038B-18	Sequence 18, Appli
8	79	7.5	1132	1	US-08-446-010B-18	Sequence 18, Appli
9	79	7.5	1132	2	US-08-805-445-18	Sequence 18, Appli
10	79	7.5	1132	2	US-08-064-067D-18	Sequence 18, Appli
11	79	7.5	1142	1	US-09-066-208-18	Sequence 18, Appli
12	79	7.5	1142	1	US-08-097-997A-11	Sequence 11, Appli
13	79	7.5	1142	4	US-08-665-574C-11	Sequence 11, Appli
	79	7.5	1142	4	US-08-346-594-11	Sequence 11, Appli
	79	7.5	1154	1	US-08-357-598-7	Sequence 7, Appli
	79	7.5	1154	1	US-08-446-010B-24	Sequence 24, Appli
	79	7.5	1154	2	US-09-003-289-7	Sequence 7, Appli
	7.5	1154	5	PCT-US95-16435-7	Sequence 7, Appli	
	7.5	547	4	US-09-877-730-26	Sequence 26, Appli	
	5	624	4	US-09-877-730-26	Sequence 24, Appli	
		628	4	US-09-877-730-26	Sequence 30, Appli	
		712	4	US-09-877-730-30	Sequence 22, Appli	
		83	4	US-09-877-730-22	Sequence 28, Appli	
		83	4	US-09-877-730-28	Sequence 16, Appli	
			4	US-09-877-730-16	Sequence 6, Appli	
				US-09-877-730-6	Sequence 20, Appli	
				US-09-877-730-20	Sequence 10, Appli	
				US-09-877-730-10		

## ALIGNMENTS

RESULT 1  
 US-08-665-040-2  
 ; Sequence 2, Application US/08665040  
 ; Patent No. 5869318  
 ; GENERAL INFORMATION:  
 APPLICANT: ESTIVILL PALLEJA, XAVIER  
 APPLICANT: FUENTES, JUAN JOSE  
 APPLICANT: PRITCHARD, MELANIE  
 TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
 TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
 TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,  
 TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (PSCR1) HIGHLY  
 TITLE OF INVENTION: EXPRESSED IN FOSTEL BRAIN AND IN HEART AND METHOD  
 TITLE OF INVENTION: FOR CHARACTERIZING IT.  
 NUMBER OF PAGES: 1

28	78.5	7.5	991	4	US-09-877-730-12	Sequence 12, Appl
29	78.5	7.5	1069	4	US-09-877-730-2	Sequence 2, Appl
30	78.5	7.5	1072	4	US-09-877-730-18	Sequence 18, Appl
31	78.5	7.5	1150	4	US-09-877-730-8	Sequence 8, Appl
32	78.5	7.5	1706	2	US-08-459-568-2	Sequence 2, Appl
33	78.5	7.5	1706	2	US-08-399-411-2	Sequence 2, Appl
34	78.5	7.5	1706	3	US-08-516-859A-2	Sequence 2, Appl
35	78.5	7.5	1706	4	US-09-586-475-2	Sequence 2, Appl
36	78.5	7.5	1706	4	US-09-528-706-2	Sequence 2, Appl
37	77	7.3	362	2	US-08-985-090-5	Sequence 5, Appl
38	77	7.3	362	3	US-09-165-543-32	Sequence 32, Appl
39	77	7.3	532	6	RE34606-6	Patent No. RE34,60
40	77	7.3	1912	3	US-08-913-832A-2	Sequence 2, Appl
41	77	7.3	1912	4	US-09-249-161A-2	Sequence 2, Appl
42	76	7.3	445	3	US-09-165-543-5	Sequence 5, Appl
43	75	7.2	1158	4	US-09-060-442-2	Sequence 2, Appl
44	73.5	7.0	560	2	US-08-928-692-53	Sequence 53, Appl
45	73.5	7.0	560	4	US-09-333-972-53	Sequence 53, Appl

Query Match	8.4%	Score 88;	DB 5;	Length 284;	
Best local Similarity	24.1%	Pred. No. 0.033;	Indels 40;	Gaps 7	
Matches	40;	Conservative 24;	Mismatches 62;		
QY	27	LCKAKESLSLFRITDKDTTEQYFKRSFKRVIRINSNLSAADARLRHKTEPLGKEMKAYFA	86		
	:				
Db	3	MCYGMGSLFRSSROTLPFA-----SQWRVAD-----LYAASFIN--LLEYYP	43		
QY	87	QTLHGSSHLAPNP-----DKQFLISPPAS---PFGMKQYEDATPVINYDLLYAI	135		
	:				
Db	44	FSYLFRAAHVLMPEHSTVEVETHVDINDIMESPLATNRRTSVDEKDK---TDYKRNDLRSI	99		

Patent No. 6420155  
 Applicant: Kerry E. Quinn  
 Applicant: Curagen Corporation  
 Title of Invention: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids  
 Title of Invention: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids  
 File Reference: 15966-581  
 Current Application Number: US/09/641,741  
 Current Filing Date: 2000-08-18  
 Prior Application Number: 60/155,613  
 Prior Filing Date: 1999-10-14  
 Prior Application Number: 60/175,534  
 Prior Filing Date: 2000-01-12  
 Prior Application Number: 60/224,086  
 Prior Filing Date: 2000-08-09  
 Number of SEQ ID NOS: 32  
 Software: PatentIn Ver. 2.0  
 SEQ ID NO 30  
 Length: 1128  
 Type: PRT  
 Organism: Mus musculus



US-09-641-341-30

Query Match  
 Best Local Similarity 7.8%; Score 82; DB 4; Length 1128;  
 Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;  
 QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93  
 Db 977 NWKRIEILAMNGNRPILGVDPSPMTPOQRMOORLQYRLRMREQML---RRLNSTA 1033  
 QY 94 SHLAPPNDKFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136  
 Db 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETTYTEVVT----- 1081  
 QY 137 KLGPEKVELHAADTPPSVVHVHVCSDQNEEEEEEME 175  
 Db 1082 -----EFTEYGT-----LEVEIEEEEEEEEMD 1108

## RESULT 5

US-09-060-482-8  
 ; Sequence 8, Application US/09060482  
 ; Patent No. 6458766  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, Mu-En  
 ; APPLICANT: Layne, Matthew D.  
 ; APPLICANT: Yee, Shaw-Fang  
 ; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE  
 ; FILE REFERENCE: 05433/036001  
 ; CURRENT APPLICATION NUMBER: US/09/060,482  
 ; EARLIER FILING DATE: 1998-04-15  
 ; EARLIER FILING DATE: 1997-03-14  
 ; EARLIER FILING DATE: 1996-03-15  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1128  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-060-482-8

Query Match  
 Best Local Similarity 7.8%; Score 82; DB 4; Length 1128;  
 Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;  
 QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93  
 Db 977 NWKRIEILAMNGNRPILGVDPSPMTPOQRMOORLQYRLRMREQML---RRLNSTA 1033  
 QY 94 SHLAPPNDKFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136  
 Db 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETTYTEVVT----- 1081  
 QY 137 KLGPEKVELHAADTPPSVVHVHVCSDQNEEEEEEME 175  
 Db 1082 -----EFTEYGT-----LEVEIEEEEEEEEMD 1108

## RESULT 6

US-08-111-939-2  
 ; Sequence 2, Application US/08111939  
 ; Patent No. 5460951  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kawai, Shinji  
 ; APPLICANT: Takeshita, Sunao  
 ; APPLICANT: Okazaki, Makoto  
 ; APPLICANT: Amann, Egon  
 ; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like  
 ; OF INVENTION: Protein and Process for its Production  
 ; SEQUENCES: 27  
 ; ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 ADDRESSEE: Dunner  
 STREET: 1300 I Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/111,939  
 FILING DATE: 26-AUG-1993  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 324033/92  
 FILING DATE: 03-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 230029/92  
 FILING DATE: 28-AUG-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Forman, David S.  
 REGISTRATION NUMBER: 33,694  
 REFERENCE/DOCKET NUMBER: 02481.1321-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4000  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1128 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-111-939-2

Query Match  
 Best Local Similarity 7.7%; Score 81; DB 1; Length 1128;  
 Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;  
 QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93  
 Db 977 NWKRIEILAMNGNRPILGVDPSPMTPOQRMOORLQYRLRMREQML---RRLNSTA 1033  
 QY 94 SHLAPPNDKFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136  
 Db 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETTYTEVVT----- 1081  
 QY 137 KLGPEKVELHAADTPPSVVHVHVCSDQNEEEEEEME 175  
 Db 1082 -----EFTEYGT-----LEVEIEEEEEEEEMD 1108

## RESULT 7

US-08-446-038B-18  
 ; Sequence 18, Application US/08446038B  
 ; Patent No. 5658791  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
 ; APPLICANT: Harpur, Ailsa  
 ; TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
 ; COMPUTER: IBM PS/2

Sun Dec 15 08:38:08 2002

us-09-782-953-3.ral

Page 4

OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446.038B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064.067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5658791man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-638-3884  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-446-038B-18

Query Match 7.5%; Score 79; DB 1; Length 1132;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;  
DB 12 ATTAGCHLDPFRVFDGLCRAKFESLFTYDKDTTFOYFKSFKRVINFSNPLSAADARLRL 71  
64 AAQACRISP-----LC-----HNLFALYDENTKLMWAPN---RTTVDDKMS-----LRL 105  
QY 72 HKTFFLCKEMKLYPAOTLHIGSSHLAPPNDKQFLISPPASPPVGW---KQVEDATPVINY 129  
DB 106 H-----YRNRFFYT-----NMHGTDNDQSVWRHSPKKQKN-GYEKKKIPDATPLDA 152  
QY 130 DLTYAISKLGPEKYEYLAATDPTPSVVHVVCSDQNEBEEBEE-----ME 175  
DB 153 SLELYLPAQG---QYDLYVKCLAP-----IRDPKTEODGHDIEHCLGMVLAISHVAMMK 204  
QY 176 RMKRPK-PKLIOTRRPEYTP 194  
DB 205 KMQLPELPKDISYKR--YTP 222

RESULT 8  
US-08-446-010B-18  
Sequence 18, Application US/08446010B  
Patent No. 5716818  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pelfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446.010B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 433  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446.038  
FILING DATE: 19-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064.067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Baer, Madeline F.  
REGISTRATION NUMBER: 36,437  
REFERENCE/DOCKET NUMBER: LUD 5244.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-638-3884  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-446-010B-18

Query Match 7.5%; Score 79; DB 1; Length 1132;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;  
DB 12 ATTAGCHLDPFRVFDGLCRAKFESLFTYDKDTTFOYFKSFKRVINFSNPLSAADARLRL 71  
64 AAQACRISP-----LC-----HNLFALYDENTKLMWAPN---RTTVDDKMS-----LRL 105  
QY 72 HKTFFLCKEMKLYPAOTLHIGSSHLAPPNDKQFLISPPASPPVGW---KQVEDATPVINY 129  
DB 106 H-----YRNRFFYT-----NMHGTDNDQSVWRHSPKKQKN-GYEKKKIPDATPLDA 152  
QY 130 DLTYAISKLGPEKYEYLAATDPTPSVVHVVCSDQNEBEEBEE-----ME 175  
DB 153 SLELYLPAQG---QYDLYVKCLAP-----IRDPKTEODGHDIEHCLGMVLAISHVAMMK 204  
QY 176 RMKRPK-PKLIOTRRPEYTP 194  
DB 205 KMQLPELPKDISYKR--YTP 222

RESULT 9  
US-08-805-445-18  
Sequence 18, Application US/08805445  
Patent No. 5821069  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pelfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2

```

; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-JUN-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-805-445-18

```

Query Match 7.5%; Score 79; DB 2; Length 1132;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

```

QY 12 ATACHLDPRVFDGLCRKAFESLFTYDKDTTFQYFKSKRVINFSNPLSADARLRL 71
DB 64 AAQACRISP-----LC-----HNLFALYDENTKLYAPN---RTITVDDKMS-----LRL 105

QY 72 HKTEFLGKMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGV--KOVEDATPVINY 129
DB 106 H-----YRMRFYFT-----NMHGTNDNEQSVWRHSPKKQKN-GYEKKKIPDTPDLLDA 152

QY 130 DLLVAISKLGPGKYEHLHAATDPTSPVVHVCSQDNEEEEE-----ME 175
DB 153 SSLLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECLGMAVLAISHYAMWK 204

QY 176 RMKRPK-PKIIOTRRPEYTP 194
DB 205 KMQLPELPKDISYKR--YIP 222

```

```

RESULT 10
US-08-064-067D-18
; Sequence 18, Application US/08064067D
; Patent No. 5852184
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2

```

```

; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/064,067D
; FILING DATE: 30-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5852184man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-064-067D-18

```

Query Match 7.5%; Score 79; DB 2; Length 1132;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

```

QY 12 ATACHLDPRVFDGLCRKAFESLFTYDKDTTFQYFKSKRVINFSNPLSADARLRL 71
DB 64 AAQACRISP-----LC-----HNLFALYDENTKLYAPN---RTITVDDKMS-----LRL 105

QY 72 HKTEFLGKMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGV--KOVEDATPVINY 129
DB 106 H-----YRMRFYFT-----NMHGTNDNEQSVWRHSPKKQKN-GYEKKKIPDTPDLLDA 152

QY 130 DLLVAISKLGPGKYEHLHAATDPTSPVVHVCSQDNEEEEE-----ME 175
DB 153 SSLLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECLGMAVLAISHYAMWK 204

QY 176 RMKRPK-PKIIOTRRPEYTP 194
DB 205 KMQLPELPKDISYKR--YIP 222

```

```

RESULT 11
US-09-066-208-18
; Sequence 18, Application US/09066208
; Patent No. 5910426
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,208
; FILING DATE:

```

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/005,445  
FILING DATE: 25-FEB-1997  
APPLICATION NUMBER: US 08/446,038  
FILING DATE: 19-MAY-1995  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-JUN-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5910426man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-688-3884  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-066-208-18

Query Match 7.5%; Score 79; DB 2; Length 1132;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHLDPRVVDGLCRAPESIFRTYDQDTFOYKSFKRVINFSNPLSADARLR 71  
DB 64 AAOACRISP-----LC-----HNLFALYDENTKLTWAPN--RTTYDDKMS-----LRL 105  
QY 72 HKTFFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPVGM--KQVEDATPVINY 129  
DB 106 H-----YRMRFYFT-----NMHGTDNDSQSVRHSPPKQKN-GYEKKKIPATPLD 152  
QY 130 DLVAISKLGGEKYEHLAATDPTPSVVHVHCSDDQNEEBEE-----ME 175  
DB 153 SLELYLFAQG---GYDLVKCLAP-----IRDPKTEODGHDIENECGLMAVLAISHVAMK 204

QY 176 RMRKPK-PKTIQTRRPEYTP 194  
DB 205 KMQLPELPKDISYKR--YIP 222

RESULT 12  
US-08-097-997A-11  
Sequence 11, Application US/08097997A  
Patent No. 5728536  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Mituhnn, Bruce A.  
APPLICANT: Ouellet, Frederick W.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,997A  
FILING DATE: 29-JULY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0370000/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1142 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-097-997A-11

Query Match 7.5%; Score 79; DB 1; Length 1142;  
Best Local Similarity 24.0%; Pred. No. 4.5;  
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHLDPRVVDGLCRAPESIFRTYDQDTFOYKSFKRVINFSNPLSADARLR 71  
DB 52 AAOACRISP-----LC-----HNLFALYDENTKLTWAPN--RTTYDDKMS-----LRL 93  
QY 72 HKTFFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPVGM--KQVEDATPVINY 129  
DB 94 H-----YRMRFYFT-----NMHGTDNDSQSVRHSPPKQKN-GYEKKKIPATPLD 140  
QY 130 DLVAISKLGGEKYEHLAATDPTPSVVHVHCSDDQNEEBEE-----ME 175  
DB 141 SLELYLFAQG---GYDLVKCLAP-----IRDPKTEODGHDIENECGLMAVLAISHVAMK 192

QY 176 RMRKPK-PKTIQTRRPEYTP 194  
DB 193 KMQLPELPKDISYKR--YIP 210

RESULT 13  
US-08-665-574C-11  
Sequence 11, Application US/08665574C  
Patent No. 6136595  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Mituhnn, Bruce A.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,574C  
FILING DATE: 18-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/282,012  
FILING DATE: 29-JUL-1994  
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/097,997
; FILING DATE: 29-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/118,968
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-574C-11

```

```

Query Match          7.5% Score 79; DB 4; Length 1142;
Best Local Similarity 24.0% Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

```

```

QY 12 ATACHLDPRVVDGLCRAPESLFTYDKDTTFQYFKSRVRINFSPLSAADARLRL 71
DB 52 AAQACRISP-----LC-----HNLFALYDENTKLVAPN---RTITVDDKMS-----LRL 93
QY 72 HKTFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 94 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKQKN-GYEKKKIPDPTLLDA 140
QY 130 DLLVAISKLGPGKGYELHAATDPTPSVVHVCSQDENEDEEE-----ME 175
DB 141 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLMAVLAISHYAMWK 192
QY 176 RMKRPK-PKIIOTRPEYTP 194
DB 193 KMQLPELPKDISYKR--YIP 210

```

```

RESULT 14
US-08-946-994-11
; Sequence 11, Application US/08946994
; Patent No. 6210654
; GENERAL INFORMATION:
; APPLICANT: Ihle, James N.
; APPLICANT: Silvenoinen, Ollie
; APPLICANT: Witthuhn, Bruce A.
; APPLICANT: Quelle, Frederick W.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,994
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,574
; FILING DATE: 18-JUN-1996
; REGISTRATION NUMBER: 08/282,012

```

```

; FILING DATE: 29-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/097,997
; FILING DATE: 29-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/118,968
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-946-994-11

```

```

Query Match          7.5% Score 79; DB 4; Length 1142;
Best Local Similarity 24.0% Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;
QY 12 ATACHLDPRVVDGLCRAPESLFTYDKDTTFQYFKSRVRINFSPLSAADARLRL 71
DB 52 AAQACRISP-----LC-----HNLFALYDENTKLVAPN---RTITVDDKMS-----LRL 93
QY 72 HKTFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 94 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKQKN-GYEKKKIPDPTLLDA 140
QY 130 DLLVAISKLGPGKGYELHAATDPTPSVVHVCSQDENEDEEE-----ME 175
DB 141 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLMAVLAISHYAMWK 192
QY 176 RMKRPK-PKIIOTRPEYTP 194
DB 193 KMQLPELPKDISYKR--YIP 210

```

```

RESULT 15
US-08-357-598-7
; Sequence 7, Application US/08357598
; Patent No. 5705625
; GENERAL INFORMATION:
; APPLICANT: Clavin, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,598
; FILING DATE: 15-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033001
; TELECOMMUNICATION INFORMATION:

```

Sun Dec 15 08:38:08 2002

us-09-782-953-3.ral

TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1154 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-357-598-7

Query Match 7.5%; Score 79; DB 1; Length 1154;  
Best Local Similarity 24.0%; Pred. No. 4.6; Mismatches 64; Indels 58; Gaps 13;  
Matches 48; Conservative 30;

OY 12 ATACHIDPRVFDGLCRAKESLFRITDQDTFOYFKSFKVRINFNSNPLSADARLRL 71  
Db 64 AAQCRISP-----LC---HNLFLYDENTKLMYAPN---RTIVDDKMS-----IRL 105  
OY 72 HKTFLGKEMKLYAQTUHGSSHLAPPNDKQFLISPPASPPVGM--KQVEDATPVINY 129  
Db 106 H-----YRMRFYFT-----NMHGTNDNEOSVWRHSPKKQKN-GYEKKKIPDATPLIDA 152  
OY 130 DLIVAIKSLGPEKYLHAATDPTPSVVVHVCSDSQENEDEEDE-----ME 175  
Db 153 SLEYLEFAOG---QYDLVKCLAP-----IRDPKTEODGDHIDENECLGMAYLAISHYAMMK 204  
OY 176 RMKRPK-PKTIQTRRPETP 194  
Db 205 KMQLPELPKDISYKR--YIP 222

Search completed: December 11, 2002, 11:39:12  
Job time: 15.3524 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 33.0556 Seconds  
(without alignments)  
2368.771 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDQLPSATACHLDP.....RPKIKIOTRRPEYPIHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPO.spool/US09782953/runat\_11122002\_11431\_17465/app.query.fasta\_1.1173  
-DB=Published Applications NA -OFFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09782953@cgn 1 1 38 @runat\_11122002\_11431\_17465  
-NCFU=6 -ICFPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US50\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query	Match	Length	ID	Description
	100.0	597	10	US-09-782-953-2	Sequence 2, Appli	
	95.7	2212	10	US-09-782-953-14	Sequence 14, Appl	
	2358	10	US-09-925-302-347	Sequence 347, App	Sequence 5, Appli	
	597	10	US-09-782-953-5			

5	872.5	83.3	2331	10	US-09-782-953-11	Sequence 11, Appl
6	872.5	83.3	2355	12	US-10-044-050-255	Sequence 255, App
7	847.5	80.9	2173	10	US-09-880-107-3340	Sequence 3340, Ap
8	749.5	71.5	599	10	US-09-782-953-1	Sequence 1, Appli
9	613.5	58.5	594	10	US-09-782-953-8	Sequence 8, Appli
10	596.5	56.9	3184	10	US-09-954-456-497	Sequence 497, App
11	596.5	56.9	3184	10	US-09-782-953-17	Sequence 17, Appl
12	596	56.9	828	10	US-09-782-953-20	Sequence 20, Appl
13	595	56.8	720	10	US-09-782-953-23	Sequence 23, Appl
14	285.5	27.2	412	10	US-09-864-761-10388	Sequence 10388, A
15	285.5	27.2	446	10	US-09-864-761-2064	Sequence 2064, Ap
16	266.5	25.4	486	10	US-09-864-761-809	Sequence 809, App
17	200	19.1	365	10	US-09-728-445-736	Sequence 736, App
18	187.5	17.9	123	10	US-09-864-761-17592	Sequence 17592, A
19	181	17.3	111	10	US-09-864-761-27019	Sequence 27019, A
20	140	13.4	85	10	US-09-864-761-18808	Sequence 18808, A
21	90.5	8.6	12308	9	US-09-854-133-422	Sequence 422, App
22	90.5	8.6	12308	10	US-09-738-973-422	Sequence 422, App
23	87	8.3	4600	9	US-09-736-457-1797	Sequence 1797, Ap
24	87	8.3	4600	9	US-09-902-941-1797	Sequence 1797, Ap
25	83.5	8.0	32191	10	US-09-764-877-3374	Sequence 3374, Ap
26	82.5	7.9	530	10	US-09-864-761-15897	Sequence 15897, A
27	81.5	7.8	2095	12	US-10-044-090-665	Sequence 665, App
28	81	7.7	3138	10	US-09-880-107-1716	Sequence 1716, Ap
29	80.5	7.7	1211	9	US-09-981-876-123	Sequence 123, App
30	80.5	7.7	6799	9	US-09-902-941-1883	Sequence 1883, Ap
31	80.5	7.7	16798	10	US-09-070-927A-195	Sequence 195, App
32	80	7.6	552	10	US-09-954-456-594	Sequence 594, App
33	79	7.5	2027	10	US-09-771-161A-29	Sequence 29, Appl
34	79	7.5	3541	10	US-09-880-107-2379	Sequence 2379, Ap
35	79	7.5	3775	10	US-09-954-456-2221	Sequence 2210, Ap
36	79	7.5	3775	10	US-09-880-107-2221	Sequence 2221, Ap
37	79	7.5	36651	10	US-09-964-469-3	Sequence 3, Appli
38	78.5	7.5	569	9	US-10-040-739-301	Sequence 301, App
39	78.5	7.5	738	10	US-09-974-300-1099	Sequence 1099, Ap
40	78.5	7.5	4108	10	US-09-883-096-1	Sequence 1, Appli
41	77	7.3	1086	9	US-09-349-755-33	Sequence 33, Appl
42	77	7.3	1086	9	US-09-166-334-33	Sequence 33, Appl
43	77	7.3	1086	10	US-09-350-206-33	Sequence 33, Appl
44	77	7.3	1440	9	US-09-712-363-94	Sequence 94, Appl
45	77	7.3	2218	9	US-09-349-755-31	Sequence 31, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-782-953-2  
; Sequence 2, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: US/09/782,953  
; CURRENT APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(594)  
US-09-782-953-2

Alignment Scores:  
Pred. No.: 2.07e-119  
Score: 1048.00  
Length: 597  
Matches: 198

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x US-09-782-953-2 (1-597)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHISLeuAspPro 20  
 DB 1 ATGAGAGAGGTGATCTGACAGACCTGCCAGACCATGCTGCTGACCTGGACCCG 60  
 QY 21 ArgValPheValAspIleuCysArgAlaIlyPheGluSerLeuPheArgThrTyArg 40  
 DB 61 CGCGTGTGTCGACGCGCTGCTCGGCGCAATTGAACTCCCTTCAGAACATATGAC 120  
 QY 41 LysAspThrThrPheGlnTyRheLeuSerPheLysArgValArgIleAsnPheSerAsn 60  
 DB 121 AAGGACACCACTTCCAGTATTTAAAGCTTCAACGATCCCTTCAGAACATATGAC 180  
 QY 61 ProLeuSerAlaIleAspAlaArgLeuArgLeuHISLysThrGluPheLeuGlyLysGlu 80  
 DB 181 CCTTATCTGACAGCCGATGCCAGGCTCGGCTGCACAAAGACGAGTTCCTGGGAAAGAA 240  
 QY 81 MetLysLeuTyRheAlaGlnThrLeuHISLysSerSerIleLeuAlaProProAsn 100  
 DB 241 ATGAAGTTATTTTGGCTCAGACTTACACATAGAAAGTTCAACCTGGCTCGCCCAAT 300  
 QY 101 ProAspLysGlnPheLeuIleSerProAlaSerProProValGlyTyrLysGlnVal 120  
 DB 301 CCCCAACAAAGTTCCTCATCTCCCTCCGCTCCTCCCTCGGCTGGAACAAAGTA 360  
 QY 121 GluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140  
 DB 361 GAAGATGCCACCCCGCATTAATTACATCTTTATATGCACTCCAAAGCTGGAGCCA 420  
 QY 141 GlyGlyLysTyRGlulLeuHISAlaIleThrAspProThrProSerValIleValHISVal 160  
 DB 421 GGAGAGAGATGAAGTCACTGATGACAGCCAGACCCCACTCCCACTGGTGTGTCACGTG 480  
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
 DB 481 TGTGAGAGTGAACCAAGAAATGAGAGAGAAAGAAAGAAATGAGAAAGAAAGAAAGAAAGCC 540  
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHISLeuSer 198  
 DB 541 AAGCCAAAATATTCACAGACGAGACCGAGTACACACGATCCACTTACG 594

RESULT 2  
 US-09-782-953-14

/ Sequence 14, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHERMEL, BEVERLY  
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 / TITLE OF INVENTION: CALCIUM INTERACTING PROTEIN (MCIP)  
 / FILE REFERENCE: UTSD-674P21  
 / CURRENT APPLICATION NUMBER: US/09/782,953  
 / CURRENT FILING DATE: 2001-02-13  
 / PRIOR APPLICATION NUMBER: 60/216,601  
 / PRIOR FILING DATE: 2000-07-07  
 / NUMBER OF SEQ ID NOS: 27  
 / SOFTWARE: Patent Ver. 2.1  
 / SEQ ID NO 14  
 / LENGTH: 2212  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (25)..(615)  
 / US-09-782-953-14

Alignment Scores:

Pred. No.: 5,266-113 Length: 2212  
 Score: 1002.50 Matches: 191  
 Percent Similarity: 97.98% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 1  
 Query Match: 95.66% Indels: 1  
 DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-782-953-14 (1-2212)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHISLeuAspPro 20  
 DB 25 ATGAGAGAGGTGATCTGACAGACCTGCCAGACCATGCTGCTGACCTGGACCCG 84  
 QY 21 ArgValPheValAspIleuCysArgAlaIlyPheGluSerLeuPheArgThrTyArg 40  
 DB 85 CGCGTGTGTCGACGCGCTGCTCGGCGCAATTGAACTCCCTTCAGAACATATGAC 144  
 QY 41 LysAspThrThrPheGlnTyRheLeuSerPheLysArgValArgIleAsnPheSerAsn 60  
 DB 145 AAGGACATCACTTCCAGTATTTAAAGCTTCAACGATCCCTTCAGAACATATGAC 204  
 QY 61 ProLeuSerAlaIleAspAlaArgLeuArgLeuHISLysThrGluPheLeuGlyLysGlu 80  
 DB 205 CCTTATCTGACAGCCGATGCCAGGCTCGGCTGCACAAAGACGAGTTCCTGGGAAAGAA 264  
 QY 81 MetLysLeuTyRheAlaGlnThrLeuHISLysSerSerIleLeuAlaProProAsn 100  
 DB 265 ATGAAGTTATTTTGGCTCAGACTTACACATAGAAAGTTCAACCTGGCTCGCCCAAT 324  
 QY 101 ProAspLysGlnPheLeuIleSerProAlaSerProProValGlyTyrLysGlnVal 120  
 DB 325 CGAGCAACAAAGTTCCTCATCTCCCTCCGCTCCTCCCTCGGCTGGAACAAAGTA 384  
 QY 121 GluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140  
 DB 385 GAAGATGCCACCCCGCATTAATTACATCTTTATATGCACTCCAAAGCTGGAGCCA 444  
 QY 141 GlyGlyLysTyRGlulLeuHISAlaIleThrAspProThrProSerValIleValHISVal 160  
 DB 445 GGGGAAAGTGAAGTTCACAGACCATGATGATCTTTATATGCACTCCCAAGCTGGTGTGTCACGTG 504  
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180  
 DB 505 TGTGAGAGTGAACCAAGAAATGAGAGAGAAAGAAAGAAATGAGAAAGAAAGAAAGAAAGCC 561  
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHISLeuSer 198  
 DB 562 AAGCCAAAATATTCACAGACGAGGCGGAGTACACGCGATCCACTTACG 615

RESULT 3

US-09-925-302-347  
 / Sequence 347, Application US/09925302  
 / Patent No. US2002004941A1  
 / GENERAL INFORMATION:  
 / APPLICANT: ROSEN ET AL.  
 / TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 / FILE REFERENCE: PA104  
 / CURRENT APPLICATION NUMBER: US/09/925,302  
 / CURRENT FILING DATE: 2001-08-10  
 / PRIOR APPLICATION NUMBER: PCT/US00/05918  
 / PRIOR FILING DATE: 2000-03-08  
 / PRIOR APPLICATION NUMBER: 60/124,270  
 / PRIOR FILING DATE: 1999-03-12  
 / NUMBER OF SEQ ID NOS: 896  
 / SOFTWARE: Patent Ver. 2.0  
 / SEQ ID NO 347  
 / LENGTH: 2358  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / US-09-925-302-347

Alignment Scores: 1,866-107 Length: 2358  
 Pred. No.: 1,866-107 Length: 2358



Score: 957.50 Matches: 188  
Percent Similarity: 96.46% Conservative: 3  
Best Local Similarity: 94.95% Mismatches: 6  
Query Match: 91.36% Indels: 2  
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-925-302-347 (1-2358)

```
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 114 ATGGAGAGTGGAGCTGCGAGACCTGCCCGGCGCCACCATCGCTGTCACCTGGACCG 173
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 174 CGCGTTCGTGGACGGCTGTGCGGGGCAAAATTTAGTCCCTCTTTAGGACGATATGAC 233
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 234 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGAAATAAATTCAGCAAC 293
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 294 CCCTTCTCCGACGACATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 353
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysThrGluPheLeuGlyLysGlu 100
Db 354 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCA-AAAT 412
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 413 CCAGCAAGCAGTTTCGATCTCCCTCCGCGCTCTCCGSCAGTGGGATGGAAACAAGTG 472
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 473 GAAGATGCGACCCAGCTCATAACTATGATCTCTTATATGCCATCTCCAGCTGGGCGCA 532
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 533 GGGGAAAAGTATGAATTCACGCGAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 592
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 593 TGTAGAGTGTATCAAGAG---AAGGAGGAAGAGAGAAATGGAAGATGAGGAGACCT 649
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 650 AAGCCAAAATTTATCCAGACGAGGAGCGCGAGTACACGCGGATCCACCTCAGC 703
```

## RESULT 4

US-09-782-953-5  
; Sequence 5, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD-674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(594)  
US-09-782-953-5

Alignment Scores:  
Pred. No.: 8,33e-104 Length: 597  
Score: 920.50 Matches: 178  
Percent Similarity: 92.42% Conservative: 5  
Best Local Similarity: 89.90% Mismatches: 12  
Query Match: 87.83% Indels: 3  
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-782-953-5 (1-597)

```
Qy 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspPro 20
Db 1 ATGATATTTAGGACTTTAGTCAATTTTAGCTCCCTGATTTGCTGTGGCAACCAT 60
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 61 GATGTCTTCAGCGAAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 121 AAGGACACCATCTTCCAGTATTTTAAGAGCTTCAACGTGTCGATTAACCTTCAGCAAC 180
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 181 CCCATTATCTGACGCGATGCCAGCTCGCGCTGCACAGACCGAGTTCTCGGGAAGGAA 240
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysThrGluPheLeuGlyLysGlu 100
Db 241 ATGAAGTTGTATTTTGCCTCAGACTTTACACATAGGAAGTTCCACACCTGGCTCCGCCCAAT 300
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 301 CCCGCAAAACAGTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAACAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 361 GAAGATGCGACCCCGCTCATAAATTCAGTCTTTTATATGCCATCTCCAGCTGGGCGCA 420
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 421 GGAGAGAAGTATGAATTCATGTCAGCGACAGACCCCACTCCAGTGTGGTGGTCCACGCTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 481 TGTAGAGTGCACCAAGAAATGAGGAGGAAGAGAGATGGAGAGATGAGAGATGAGAGACCC 540
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 541 AAGCCAAAATTTATCCAGACGAGGAGCGGAGTACACACCGATCCACCTTAGC 594
```

## RESULT 5

US-09-782-953-11  
; Sequence 11, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD-674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (144)...(734)  
;

US-09-782-953-11

## Alignment Scores:

Pred. No.:	4,61e-97	Length:	2331
Score:	872.50	Matches:	168
Percent Similarity:	93.09%	Conservative:	7
Best Local Similarity:	89.36%	Mismatches:	12
Query Match:	83.25%	Indels:	1
DB:	10	Gaps:	1

US-09-782-953-3 (1-198) x US-09-782-953-11 (1-2331)

```

QY 11 SerAlaThrIleAlaCysHisIleuAspProArgValPheValAspGlyLeuCysArgAla 30
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 174 AGTCCCTGATTCCTGCTGTGGCAAAAGTATCTTCAGCGAAAGTGAACCGAGGCC 233
QY 31 LysPheGluSerIleuPheArgThrTyraAspLysAspThrThiPheGlnTyrPheLysSer 50
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 234 AAATTGAGTCCCTCTTAGACATGACAGACATCACCTTTTCAGTATTTTAAAGAC 293
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 294 TTCAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGCTCCAG 353
QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHis 90
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 354 CTGCATAGAGCTGAGTTCTGGGAAAGAAATGAAGTATATTTGCTCAGACCTTACAC 413
QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 414 ATGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGACGTTCTGATCTCCCTCCC 473
QY 111 AlaSerProProValGlyTyrrPheGlnValGluAspAlaThrProValIleAsnTyra 130
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 474 GCCTCTCCGCGAGTGGAGTGAACAAGTGAAGATCGACCCGACATTAACATATGAT 533
QY 131 LeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAlaThr 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 534 CTTTATATGCCATCTCCAAAGCTGGGCCAGGGGAAAGTAAATTTGACGACGAGCT 593
QY 151 AspProThrProSerValValIleHisValCysGluSerAspGlnGluGlnGluGlu 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 594 GACACACCTCCAGCGGTGGTGCATGATGAGAGTATCAAGG---AAGAGGAA 650
QY 171 GluGluGluMetLysArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 651 GAAGAGGAAATGAAAGAAATGAGAGACCTTAAGCCAAAATTATCCAGACGAGAGGCCG 710
QY 191 GluTyrrThrProIleHisLeuSer 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 711 GAGTACACGCCGATCCACCTCAGC 734

```

RESULT 6

US-10-044-090-255

Sequence 255, Application US/10044090

Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman

TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

FILE REFERENCE: PA-0028 US

CURRENT APPLICATION NUMBER: US/10/044,090

CURRENT FILING DATE: 2002-01-09

NUMBER OF SEQ ID NOS: 850

SOFTWARE: PERL Program

SEQ ID NO 255

LENGTH: 2355

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 551189CB1

US-10-044-090-255

## Alignment Scores:

Pred. No.:	4,68e-97	Length:	2355
Score:	872.50	Matches:	168
Percent Similarity:	93.09%	Conservative:	7
Best Local Similarity:	89.36%	Mismatches:	12
Query Match:	83.25%	Indels:	1
DB:	12	Gaps:	1

US-09-782-953-3 (1-198) x US-10-044-090-255 (1-2355)

```

QY 11 SerAlaThrIleAlaCysHisIleuAspProArgValPheValAspGlyLeuCysArgAla 30
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 AGTCCCTGATTCCTGCTGTGGCAAAAGTATCTTCAGCGAAAGTGAACCGAGGCC 240
QY 31 LysPheGluSerIleuPheArgThrTyraAspLysAspThrThiPheGlnTyrPheLysSer 50
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 AAATTGAGTCCCTCTTAGACATGACAGACATCACCTTTTCAGTATTTTAAAGAC 300
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 TTCAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGCTCCAG 360
QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHis 90
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 CTGCATAGAGCTGAGTTCTGGGAAAGAAATGAAGTATATTTGCTCAGACCTTACAC 420
QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 421 ATGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGACGTTCTGATCTCCCTCCC 480
QY 111 AlaSerProProValGlyTyrrPheGlnValGluAspAlaThrProValIleAsnTyra 130
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 481 GCCTCTCCGCGAGTGGAGTGAACAAGTGAAGATCGACCCGACATTAACATATGAT 540
QY 131 LeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAlaThr 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 541 CTTTATATGCCATCTCCAAAGCTGGGCCAGGGGAAAGTAAATTTGATTCACGACGAGCT 600
QY 151 AspProThrProSerValValIleHisValCysGluSerAspGlnGluGlnGluGlu 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 601 GACACACCTCCAGCGGTGGTGCATGATGAGAGTATCAAGG---AAGAGGAA 657
QY 171 GluGluGluMetLysArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 658 GAAGAGGAAATGAAAGAAATGAGAGACCTTAAGCCAAAATTATCCAGACGAGAGGCCG 717
QY 191 GluTyrrThrProIleHisLeuSer 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 718 GAGTACACGCCGATCCACCTCAGC 741

```

RESULT 7

US-09-880-107-3340

Sequence 3340, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 3340

LENGTH: 2173

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 551189CB1

US-10-044-090-255

FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833  
US-09-880-107-3340

Alignment Scores:  
Pred. No.: 4,77e-94 Length: 2173  
Score: 847.50 Matches: 162  
Percent Similarity: 97.63% Conservative: 3  
Best Local Similarity: 95.86% Mismatches: 3  
Query Match: 80.87% Indels: 1  
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-880-107-3340 (1-2173)

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49  
DB 58 GCCAAATTTGAGTCCCTCTTTAGGACGATGACAAGGACATCACCTTTCAGTATTTAAG 117  
QY 50 SerPheLysArgValArgLeuAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
DB 118 AGCTTCAACGAGTCAGAAATAAATTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177  
QY 70 ArgLeuHisLysThrGluPheGluGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89  
DB 178 CAGCTGCATAAGACTGAGTTTCTGGAAAGGAATGAGTTTATTTTGTCTCAGACCTTA 237  
QY 90 HisLeuGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109  
DB 238 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCT 297  
QY 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyr 129  
DB 298 CCCGCTCTCCGCGAGTGGGATGGAACAAGTGGAAATGCAATGATGATGCGACGCG 357  
QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149  
DB 358 GATCTCTATATGCACTCTCCAGCTGGGCCAGGGGAAGAATGATGATGCGACGCG 417  
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
DB 418 ACTGACACCACTCCAGCGTGGTGTCCATGTATGTGAGAGTGCATCAAGAG---AAGGAG 474  
QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysLeuIleGlnThrArgArg 189  
DB 475 GAAGAAGGGAATGGAAAGATGAGGAGACCTTAAGCAAAAATTTATCCAGACGAGG 534  
QY 190 ProGluTyrThrProLeuHisLeuSer 198  
DB 535 CCGGAGTACACGCGGATCCACCTCAGC 561

## RESULT 8

US-09-782-953-1  
; Sequence 1, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-782-953-1

## Alignment Scores:

Pred. No.: 3.07e-66 Length: 594  
Score: 613.50 Matches: 119

Pred. No.: 7.07e-83 Length: 599  
Score: 749.50 Matches: 151  
Percent Similarity: 90.12% Conservative: 4  
Best Local Similarity: 87.79% Mismatches: 14  
Query Match: 71.52% Indels: 4  
DB: 10 Gaps: 1  
US-09-782-953-3 (1-198) x US-09-782-953-1 (1-599)  
QY 5 AspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspProArg 21  
DB 83 GATTTTAGGACTTTAGCTACAATTTAGCTCCCTCATTTGTTGTGGCAACGATGAT 142  
QY 22 ValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAspLys 41  
DB 143 GTCTTCAGGAAAGTAGTGAGACCGGGCCAAATTTGAATCCCTCTTCAGAACATATGACAAG 202  
QY 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61  
DB 203 GACACCACTTCCAGTATTTTAAGAGCTTCAACGTTGTCGGATTAACCTTCAGCAACCC 262  
QY 62 LeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGluMet 81  
DB 263 TTATCTGCAGCCGATGCCAGGCTGGCGCTGCACAAGACCGAGTTCTCTGGGAAGGAAATG 322  
QY 82 LysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsnPro 101  
DB 323 AGTTGTATTTTGTCTCAGACTTTACATAGGAAGTTTACACCTCGCTCCGCAAT-CCC 381  
QY 102 AspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnValGlu 121  
DB 382 GACAAACAGTCTCATCTCCCTCCGCTCTCTCTCGTGGCTGGAAACAGTAGAA 441  
QY 122 AspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGly 141  
DB 442 GATGCCACCCCGTCATAAATTAACGATCTTTATATGCAATCTCCAAGCTGGGGCCAGGA 501  
QY 142 GluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisValCys 161  
DB 502 GAGAAGTAGTACTGCATGCGACGACAGACCACTCCAGTGTGGTCCACGTGTGT 561  
QY 162 GluSerAspGlnGluAsnGluGluGluGlu 173  
DB 562 CAGAGTGAACCAAGAGATGAGGAGGAAGAGAG 597

## RESULT 9

US-09-782-953-8  
; Sequence 8, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 8  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(591)  
US-09-782-953-8

## Alignment Scores:

Pred. No.: 3.07e-66 Length: 594  
Score: 613.50 Matches: 119



```

; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(780)
US-09-782-953-17

```

```

Alignment Scores:
Pred. No.: 4,36e-63 Length: 3184
Score: 596.50 Matches: 117
Percent Similarity: 73.40% Conservative: 21
Best Local Similarity: 62.23% Mismatches: 41
Query Match: 56.92% Indels: 9
DB: 10 Gaps: 2

```

US-09-782-953-3 (1-198) x US-09-782-953-17 (1-3184)

```

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValaspGlyLeuCysArgAla 30
Db 220 TCACCTGTTGCTGCTGGTGGATGTCAGGCTCTTTACCAATCAGGAGGTTAAGGAA 279
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
Db 280 AATTTGGGGGACTGTTTCGACTTATGACTGTGTGACGTTCCAGCTATTTAAGATT 339
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 340 TTCAGAGCTGCTCGGTATAAATTCAGCAATCTTAAATCTGCAGCCGAGCTAGGATAGAG 399
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 400 CTTTCATGAACCCAAATTCAGAGGAAATAATTAAGCTCTACTTTGCACAGGTTTCAGACT 459
Qy 91 IleGlySer-----HisLeuAlaProAsnProAspLysGlnPheLeu 106
Db 460 CCAGACAGACAGATGAGACAAATGCACTGGCTCCACCCAGGCTCCCAACAGTTTCTC 519
Qy 107 IleSerProProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProVal 126
Db 520 ATCTCCGCTCTTCTCTCCACCTGTTAGCTGGCAGCCCATCAACGATGCCAGCCAGTC 579
Qy 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
Db 580 CTCNACTATGACTCTCTATGCTGTGGCCAACTAGGACCAAGAGAGATATGAGCTC 639
Qy 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
Db 640 CATGCAGGACTGAGTCCACCCCACTGCTGTCGACGTGTGCGACGTGCGACGTGACATAGAG 699
Qy 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleGln 186
Db 700 GAAGAAGAGGAGGCAAAAG-----ACTTCCCAAGGCAAAATCATCCAA 744
Qy 187 ThrArgArgProGluTyrThrPro 194
Db 745 ACTCGGCTCTGGCTGCACCC 768

```

RESULT 12

```

US-09-782-953-20
; Sequence 20, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

```

```

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)..(745)
US-09-782-953-20

```

```

Alignment Scores:
Pred. No.: 6,94e-64 Length: 828
Score: 596.00 Matches: 120
Percent Similarity: 72.16% Conservative: 20
Best Local Similarity: 61.86% Mismatches: 44
Query Match: 56.87% Indels: 10
DB: 10 Gaps: 2

```

US-09-782-953-3 (1-198) x US-09-782-953-20 (1-828)

```

Qy 2 GluGluValAspLeuGluAspLeuProSerAlaThrIleAlaCysHisLeuAspProArg 21
Db 134 GAGATGATGGATTTAAAGTATCTGCTACCTACCTTTTGTTCAGCGCTCCATGAAGCA 193
Qy 22 ValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAspLys 41
Db 194 GTGTTTGGAGCCAGCAGACAGAGAAAGATTGAAGCACTCTTCACCATCTATGATGAC 253
Qy 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61
Db 254 CAGGTTACTTTTTCAGCTGTTTAAAGCTTTAGAAAGAGTCAGATAAATTTTCAGCAAACT 313
Qy 62 LeuSerAlaAlaAspAlaArgLeuArgLysHisLysThrGluPheLeuGlyLysGluMet 81
Db 314 GAAGCGGCGAGCAGAGCGCGAATAGAACTCCAGAAAACAGACTTCAATGGCGCAGAAGCTA 373
Qy 82 LysLeuTyrPheAlaGlnThrLeuHisIleGly-----SerSerHisLeuAla 97
Db 374 AAGCTATATTTTGCACAGGTCAGATGTCGCGCGAAGTGGCGGCAAGTCTATCTCTG 433
Qy 98 ProProAsnProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyr 117
Db 434 CCGCCCGAGCCTGTCAAGCAGTCTCTCATCTCCCTCCAGCTCTCCCGCAGTGGGTGG 493
Qy 118 LysGlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLys 137
Db 494 AAGCAGAGCGAAGATGCGATGCTGTTATATAATATGATTTACTCTGTGTGTTTCCAAA 553
Qy 138 LeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValVal 157
Db 554 TTGGGACCGAGGAGAAATATGAATTCACCGCGGGAACAGAGTCGACCCAGCGTGTGTG 613
Qy 158 ValHisValCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMet 177
Db 614 GTTCATGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 655
Qy 178 LysArgProLysProLysIleGlnThrArgArgProGlu 191
Db 556 AAAAAACCCCAACAGAAAAATTGCCACAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 697

```

RESULT 13

```

US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS

```



QY 36 AlalysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49  
Db 410 GCCAAATTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTCAGTATTTTAAG 351  
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
Db 350 AGCTTCAACAGAGTCAGAAATAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 291  
QY 70 ArgLeuHisLysThrGluPheLeuGluGlyMetLysLeuTyrPheAlaGlnThrLeu 89  
Db 290 CAGCTGCATAAGAGTGGTCTTGGGAAAGGAAATGAAGTATATTTGCTCAGGTGAGT 231  
QY 90 HisIle-----GlySerSerHisLeuAlaProProAsn 100  
Db 230 TGGGTTTCATTGCTATGATGCTCTTCTCCCTCTCCCTCCCTCC 189

## RESULT 15

US-09-864-761-2064/c  
Sequence 2064, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeonica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2064  
LENGTH: 446  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: MAP TO AP000122.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4

US-09-864-761-2064

## Alignment Scores:

Pred. No.: 2,79e-26 Length: 446  
Score: 285.50 Matches: 58  
Percent Similarity: 82.43% Conservative: 3  
Best Local Similarity: 78.38% Mismatches: 10  
Query Match: 27.24% Indels: 3  
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-864-761-2064 (1-446)

QY 30 AlalysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49  
Db 444 GCCAAATTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTCAGTATTTTAAG 385  
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
Db 384 AGCTTCAACAGAGTCAGAAATAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 325  
QY 70 ArgLeuHisLysThrGluPheLeuGluGlyMetLysLeuTyrPheAlaGlnThrLeu 89  
Db 324 CAGCTGCATAAGAGTGGTCTTGGGAAAGGAAATGAAGTATATTTGCTCAGGTGAGT 265  
QY 90 HisIle-----GlySerSerHisLeuAlaProProAsn 100  
Db 264 TGGGTTTCATTGCTATGATGCTCTTCTCCCTCTCCCTCCCTCC 223

Search completed: December 15, 2002, 00:06:20

Job time : 38.0556 secs

**THIS PAGE BLANK (USPTO)**



```

; TITLE OF INVENTION: Variants of alternative splicing
;
; FILE REFERENCE: 129181.4 Compugen
;
; CURRENT APPLICATION NUMBER: US/09/724.676A
;
; CURRENT FILING DATE: 2000-11-28
;
; NUMBER OF SEQ ID NOS: 97222
;
; SOFTWARE: Patentin version 3.2
;
; SEQ ID NO 61358
;
; LENGTH: 197
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
; US-09-724-676A-61358

```

[illegible]

```

RESULT 7
US-09-724-676A-61359
; Sequence 61359, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen Ltd
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09724,676A
; *CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61359

```

```

Query Match      83.3%; Score 872.5; DB 5; Length 197;
Beat Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

11 SATTACHLDPVFVDGLCRKFESLFTYDKDITFOYKSFKRVIRNFNSPLSADABLR 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

11 SSLIACVANSIDFSESTRKFESLFTYDKDITFOYKSFKRVIRNFNSPLSADABLR 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

71 LHKTEFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQVEDATPVINYD 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

71 LHKTEFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQVEDATPVINYD 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

131 LLYATISLGPCEKELYELHAADTPSPVVVHVCSDDQNEEEEEEMERMKRPPKIIQTRRP 190
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

131 LLYATISLGPCEKELYELHAADTPSPVVVHVCSDDQ- KSEBEMERMRPKIIQTRRP 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

191 EYTPHLS 198
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

190 EYTPHLS 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 8  
US-09-724-676A-61367

```

; Sequence 61367, Application US/09724676A
;
; GENERAL INFORMATION:
;
; APPLICANT: CompuGen LTD
;
; TITLE OF INVENTION: Variants of alternative splicing
;
; FILE REFERENCE: 123181.4 CompuGen
;
; CURRENT APPLICATION NUMBER: US/09/724, 676A
;
; CURRENT FILING DATE: 2000-11-28
;
; NUMBER OF SEQ ID NOS: 97222
;
; SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 61367
;
; LENGTH: 197
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-09-724-676A-61367

```

Query Match	83.3%;	Score 872.5;	DB 5;	Length 197;
Best Local Similarity	89.4%;	Pred. No. 1.9e-81;		
Matches 168;	Conservative 7;	Mismatches 12;	Indels 1;	Gaps 1;
Qy 11	SATIACHLDPRVFDGLCRAKFESLFRTYDKDTTFQYFKGFRVRINFSPNPLSAAADARLR	70		
Db	. . . . .			
Qy 11	SSLIACVANDIIFSESTRAKFESLFRTYDKDTTFQYFKGFRVRINFSPNPLSAAADARLQ	70		
Db	. . . . .			
Qy 71	LHKTFFLGKEMKLYFAQTLHGSSHLAPPNDPKOFLISPPASPPVGWKQVEDATFVINYD	130		
Db	. . . . .			
Qy 71	LHKTFFLGKEMKLYFAQTLHGSSHLAPPNDPKOFLISPPASPPVGWKQVEDATFVINYD	130		
Db	. . . . .			
Qy 131	LLVAISKLGPEKGYELHAATDPTPSVVVHVHCESDOENEESEEMERMKRPKPKIIQTRRP	190		
Db	. . . . .			
Qy 131	LLVAISKLGPEKGYELHAATDPTPSVVVHVHCESDOE-KEEEEMERNRRPKPKIIQTRRP	189		
Db	. . . . .			
Qy 191	EYTPHLS 198			
Db 190	EYTPHLS 197			

```

RESULT 9
US-09-724-676-61356
; Sequence 61356, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61356
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Hmo sapiens
US-09-724-676-61356

```

[illegible]

Query Match	Best Local Similarity	Score	DB 5:	Length
Match 162, Conservative	95.94	847.5	DB 5:	171
			Pred. No. 5.6e-79	
			Mismatches 3; Indels 1; Gaps 1	
QY	30	AKESLFRITDKDTTFQYFKSFKRVIRINFSNPLSADARLRHKTETLGEMKLYPAQTL	89	
DB	4	AKESLFRITDKDITPQYFKSFKRVIRINFSNPLSADARLRHKTETLGEMKLYPAQTL	63	
QY	90	HIGSSHLAPNPDKQFLISPPASPPVMKQVEDATPIINDLLYAIKLGGEKTELHAA	149	
DB	64	HIGSSHLAPNPDKQFLISPPASPPVMKQVEDATPIINDLLYAIKLGGEKTELHAA	123	
QY	150	TDTPSPVVVHVHVCSDQDENEENEENEEMEMKKPKKIIOTRRPEYTPILHS	198	
DB	124	TDTPSPVVVHVHVCSDQDENEENEENEEMEMKKPKKIIOTRRPEYTPILHS	171	

Query Match	69.8%	Score 731	DB 5	Length 154
Best Local Similarity	95.2%	Pred No. 3	7e-67	
Matches 138	Conservative 3	Mismatches 4	Indels 0	Gaps 0

  

QY	1	MEEVDDLODPSATIIACHLDPVYFVDSLCRAKESLERTYDKDTTFQYFKSFKKVRINFSN	60
DB	1	MEEVDDLEDDPSATIIACHLDPVYFVDSLCRAKESLERTYDKDTTFQYFKSFKKVRINFSN	60
QY	61	PUSAADARLRLLHNTPELGKEMKLYFQOTLIHGSSHLAPRPDKQFLISPPASPPVGMKQV	120
DB	61	PFSAADARLRQLHNTPELGKEMKLYFQOTLIHGSSHLAPRPDKQFLISPPASPPVGMKQV	120
QY	121	EDATPVINVDLLVAISLGGGEKYE	145
DB	121	EDATPVINVDLLVAISLGGDWLE	145

RESULT 14  
US-09-724-676A-61353  
; Sequence 61353, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Comphen LTD

;; TITLE OF INVENTION: Variants of alternative splicing  
;; FILE REFERENCE: 129181.4 Compugen  
;; CURRENT APPLICATION NUMBER: US/09/724,676A  
;; CURRENT FILING DATE: 2000-11-28  
;; NUMBER OF SEQ ID NOS: 97222  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 61353  
;; LENGTH: 154  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-724-676A-61353

Query Match 69.8%; Score 731; DB 5; Length 154;  
Best Local Similarity 95.2%; Pred. No. 3.7e-67;  
Matches 138; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MEEVDLQDLP SATIACHLDRPVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSN 60  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 1 MEEVDLEDLP SATIACHLDRPVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSN 60  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 PLSAADARLRLHKT EFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQV 120  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 PFSADARLQLHKT EFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQV 120  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 121 EDATPVINYDLLVAISKLGGEKYE 145  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 121 EDATPVINYDLLVAISKLGPDWLE 145

RESULT 15  
US-10-290-438-2  
;; Sequence 2, Application US/10290438  
;; GENERAL INFORMATION:  
;; APPLICANT: Loring, Jeanne F.  
;; APPLICANT: Tingley, Debora W.  
;; APPLICANT: Edwards, Carla M.  
;; APPLICANT: Streeter, David G.  
;; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN  
;; FILE REFERENCE: FC-0013-1C1P  
;; CURRENT APPLICATION NUMBER: US/10/290,438  
;; CURRENT FILING DATE: 2002-11-06  
;; PRIOR APPLICATION NUMBER: 09/614,474  
;; PRIOR FILING DATE: 2000-07-11  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 2  
;; LENGTH: 255  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: Incyte ID No: 247500.5  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: 11  
;; OTHER INFORMATION: unknown or other  
US-10-290-438-2

Query Match 59.2%; Score 620.5; DB 6; Length 255;  
Best Local Similarity 62.4%; Pred. No. 1.3e-55;  
Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;  
QY 5 DLQDLP SATIACHLDRPVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSNPLSA 64  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 63 DFNDLPNSLFA CNVHQSVFEGESKEFGLFRDYDDCVTFOLF KSFRRVIRNFSNPKSA 122  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 65 ADARLRLHKT EFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVGWKQV 120  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 123 ARARLHKT EFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVGWKQV 182  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 121 EDATPVINYDLLVAISKLGGEKYEELHAATDTPSVVHVCSDOENEEEMERMKRP 180  
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 183 NDATPVINYDLLVAISKLGGEKYEELHACTESTPSVVHVCSDIIEEDPK-----TSP 237

QY 181 KPKIIOTRRPEYTP 194  
Db 238 KPKIIOTRRPGLPP 251  
Search completed: December 11, 2002, 11:50:14  
Job time : 12.6863 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:49 ; Search time 6.34401 Seconds  
(without alignments)  
506.931 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDLQDLSATIACHLDP.....RPKIIQTRPEYTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_PEP.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB\_PEP.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_PEP.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_PEP.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB\_PEP.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_PEP.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	198	10	US-09-782-953-3
2	1048	100.0	198	10	US-09-782-953-4
3	1002.5	95.7	197	10	US-09-782-953-15
4	1002.5	95.7	197	10	US-09-782-953-16
5	920.5	87.8	198	10	US-09-782-953-6
6	920.5	87.8	198	10	US-09-782-953-7
7	872.5	83.3	197	10	US-09-782-953-12
8	872.5	83.3	197	10	US-09-782-953-13
9	613.5	58.5	197	10	US-09-782-953-9
10	613.5	58.5	197	10	US-09-782-953-10
11	536.5	56.9	192	10	US-09-782-953-18
12	536.5	56.9	192	10	US-09-782-953-19
13	596	56.9	241	10	US-09-782-953-21
14	596	56.9	241	10	US-09-782-953-22
15	595	56.8	212	10	US-09-782-953-24
16	595	56.8	212	10	US-09-782-953-25
17	497	47.4	142	10	US-09-925-302-790
18	284	27.1	58	10	US-09-864-761-35379
19	284	27.1	58	10	US-09-864-761-43076

20	271.5	25.9	56	10	US-09-864-761-34111	Sequence 34111, A
21	90.5	8.6	4019	10	US-09-738-973-425	Sequence 425, App
22	86.5	8.3	822	10	US-09-824-734-3	Sequence 3, Appli
23	79	7.5	572	10	US-09-771-161A-120	Sequence 120, App
24	79	7.5	1142	10	US-09-771-161A-211	Sequence 211, App
25	78.5	7.5	864	10	US-09-883-096-2	Sequence 2, Appli
26	77	7.3	362	9	US-09-349-755-32	Sequence 32, Appli
27	77	7.3	362	9	US-09-166-334-32	Sequence 32, Appli
28	77	7.3	362	10	US-09-350-206-32	Sequence 32, Appli
29	76	7.3	445	9	US-09-349-755-5	Sequence 5, Appli
30	76	7.3	445	9	US-09-166-334-5	Sequence 5, Appli
31	76	7.3	445	10	US-09-350-206-5	Sequence 5, Appli
32	75	7.2	195	9	US-09-902-941-1888	Sequence 1888, Ap
33	75	7.2	1885	10	US-09-920-346-2	Sequence 2, Appli
34	74	7.1	311	10	US-09-819-252-2	Sequence 2, Appli
35	74	7.1	311	10	US-09-922-217-1061	Sequence 1061, Ap
36	74	7.1	311	10	US-09-833-263-1061	Sequence 1061, Ap
37	73.5	7.0	536	10	US-09-908-805B-61	Sequence 61, Appl
38	72.5	6.9	536	10	US-09-908-805B-18	Sequence 18, Appl
39	72.5	6.9	559	10	US-09-854-549-7	Sequence 7, Appli
40	72.5	6.9	559	10	US-09-836-561-6	Sequence 6, Appli
41	72.5	6.9	1924	9	US-09-866-557A-2	Sequence 2, Appli
42	72	6.9	591	10	US-09-887-586A-24	Sequence 24, Appl
43	72	6.9	591	10	US-09-903-012-24	Sequence 24, Appl
44	71.5	6.8	284	10	US-09-764-864-843	Sequence 843, App
45	71.5	6.8	590	10	US-09-815-242-10812	Sequence 10812, A

#### ALIGNMENTS

#### RESULT 1

US-09-782-953-3  
; Sequence 3 Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-782-953-3

Query Match	100.0%	Score	1048	DB	10	Length	198
Best Local Similarity	100.0%	Pred. No.	3.5e-99				
Matches	198	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MEEVDLQDLSATIACHLDP	RVFV	DGL	CR	AK	FESL
Db	1	MEEVDLQDLSATIACHLDP	RVFV	DGL	CR	AK	FESL
Qy	61	PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGMKQV	120				
Db	61	PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGMKQV	120				
Qy	121	EDATPVINYOLLVIAIKLGGKELVHAADPTSPVVHVHVCSDQNEEEEEEMERKRP	180				
Db	121	EDATPVINYOLLVIAIKLGGKELVHAADPTSPVVHVHVCSDQNEEEEEEMERKRP	180				
Qy	181	KPKIIQTRPEYTPIHLS	198				
Db	181	KPKIIQTRPEYTPIHLS	198				

```
RESULT 2
US-09-782-953-4
; Sequence 4, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-4

Query Match
Best Local Similarity 100.0%; Score 1048; DB 10; Length 198;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
QY 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
DB 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
QY 181 KPKITQTRRPETPIHLS 198
DB 181 KPKITQTRRPETPIHLS 198

RESULT 3
US-09-782-953-15
; Sequence 15, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

Query Match
Best Local Similarity 95.7%; Score 1002.5; DB 10; Length 197;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
QY 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
DB 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
QY 181 KPKITQTRRPETPIHLS 198
DB 181 KPKITQTRRPETPIHLS 197

RESULT 4
US-09-782-953-16
; Sequence 16, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match
Best Local Similarity 95.7%; Score 1002.5; DB 10; Length 197;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPVGMKQV 120
QY 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
DB 121 EDATPVINYDLVLAISKLGEGEKYELHAATDPTPSVVVHVCSDDQENEEREMERMRKP 180
QY 181 KPKITQTRRPETPIHLS 198
DB 181 KPKITQTRRPETPIHLS 197

RESULT 5
US-09-782-953-6
; Sequence 6, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
```

```
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPVFDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRPDSYNFSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-782-953-7
; Sequence 7, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPVFDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRPDSYNFSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-09-782-953-12
; Sequence 12, Application US/09782953
```

```
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-12

Query Match      83.3%; Score 872.5; DB 10; Length 197;
Best Local Similarity 89.4%; Pred. No. 2.3e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 11 SATIACHLDPVFDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSAADARL 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 SSLIACVANSIDFSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSAADARL 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 131 LLYAISKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKPKKIIOTRRP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 LLYAISKLGPGKEKYLHAATDTPSVVVHVCSDDQNEEEEMERMKPKKIIOTRRP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 191 EYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 EYTPHLS 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-782-953-13
; Sequence 13, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match      83.3%; Score 872.5; DB 10; Length 197;
Best Local Similarity 89.4%; Pred. No. 2.3e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 11 SATIACHLDPVFDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSAADARL 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 SSLIACVANSIDFSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSAADARL 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 131 LLYAISKLGPEGEKYEHLAATDPTPSVVVHVCSQDNEEBEEMERMKRPKXIIQTRRP 190  
 DB 131 LLYAISKLGPEGEKYEHLAATDPTPSVVVHVCSQDNEEBEEMERMKRPKXIIQTRRP 189  
 QY 191 EYTPHLS 198  
 DB 190 EYTPHLS 197

RESULT 9  
 US-09-782-953-9  
 ; Sequence 9, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-782-953-9

Query Match 58.5%; Score 613.5; DB 10; Length 197;  
 Best Local Similarity 63.3%; Pred. No. 4.5e-55;  
 Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

QY 11 SATIACHDPRVFDGLCRAKFESLFRTYDKDTTFQYFKSPKRVIRINSPPLSADARLR 70  
 DB 11 STLVACVVDVEVFTNOEVEKEFGELFRTYDCVTFQLFKSPFRVIRINSPKSAARARLE 70  
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126  
 DB 71 LHETQPRGKKLKYFAQVOTPETDQKHLAPQPAKQFLISPPSPVGMKQFISDATTY 130  
 QY 127 INVDLIYAISKLGPEGEKYEHLAATDPTPSVVVHVCSQDNEEBEEMERMKRPKXIIQ 186  
 DB 131 LNYDLIYAIAKLGPEGEKYEHLAAGTSTPSVVVHVCDSDMEEBEDPK-----TSPKXIIQ 185  
 QY 187 TRPEYTP 194  
 DB 186 TRRPGLP 193

RESULT 10  
 US-09-782-953-10  
 ; Sequence 10, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus

US-09-782-953-10

Query Match 58.5%; Score 613.5; DB 10; Length 197;  
 Best Local Similarity 63.3%; Pred. No. 4.5e-55;  
 Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

QY 11 SATIACHDPRVFDGLCRAKFESLFRTYDKDTTFQYFKSPKRVIRINSPPLSADARLR 70  
 DB 11 STLVACVVDVEVFTNOEVEKEFGELFRTYDCVTFQLFKSPFRVIRINSPKSAARARLE 70  
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126  
 DB 71 LHETQPRGKKLKYFAQVOTPETDQKHLAPQPAKQFLISPPSPVGMKQFISDATTY 130  
 QY 127 INVDLIYAISKLGPEGEKYEHLAATDPTPSVVVHVCSQDNEEBEEMERMKRPKXIIQ 186  
 DB 131 LNYDLIYAIAKLGPEGEKYEHLAAGTSTPSVVVHVCDSDMEEBEDPK-----TSPKXIIQ 185  
 QY 187 TRPEYTP 194  
 DB 186 TRRPGLP 193

RESULT 11  
 US-09-782-953-18  
 ; Sequence 18, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 18  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-782-953-18

Query Match 56.9%; Score 596.5; DB 10; Length 192;  
 Best Local Similarity 62.2%; Pred. No. 2.3e-53;  
 Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

QY 11 SATIACHDPRVFDGLCRAKFESLFRTYDKDTTFQYFKSPKRVIRINSPPLSADARLR 70  
 DB 6 STLVACVVDVEVFTNOEVEKEFGELFRTYDCVTFQLFKSPFRVIRINSPKSAARARLE 65  
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126  
 DB 66 LHETQPRGKKLKYFAQVOTPETDQKHLAPQPAKQFLISPPSPVGMKQFISDATTY 125  
 QY 127 INVDLIYAISKLGPEGEKYEHLAATDPTPSVVVHVCSQDNEEBEEMERMKRPKXIIQ 186  
 DB 126 LNYDLIYAIAKLGPEGEKYEHLAAGTSTPSVVVHVCDSDIEEBEDPK-----TSPKXIIQ 180  
 QY 187 TRPEYTP 194  
 DB 181 TRRPGLP 188

RESULT 12  
 US-09-782-953-19  
 ; Sequence 19, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY



```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match          56.9%; Score 596.5; DB 10; Length 192;
Best Local Similarity 62.2%; Pred. No. 2.3e-53;
Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

QY 11 SATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFNSPLSAADARLR 70
DB 11 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
6 STLVAQVVDVEFTNOEVEKFGGLFRTYDDCVTFQLFKSFRRVRINFNSPKSAAARIE 65
QY 71 LHKTEFLGKEMKLYFAQTLHGSS---HLAPPNDKQFLISPPASPPVCKQVEDATPV 126
DB 71 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
66 LHETQFGKGLKLYFAQVOTPETDQKHLHAPPQPAKQFLISPPSPPPVSWQPINATPV 125
QY 127 INVDLLYAKLGPGEKYLHAATDTPSVVVHVCSDOENEEEMERMKRKPPIQ 186
DB 127 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
126 LNYDLLYAVAKLGPGEKYLHAAGTSTPSPVVHVCDSDIEEDPK-----TSPKPKIIQ 180
QY 187 TRPEYTP 194
DB 187 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
181 TRPGLPP 188

RESULT 13
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFNSP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 ENWDLSDLTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFNSKP 97
QY 62 LSAADARLRHLKTEFLGKEMKLYFAQTLHG---SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINVDLLYAKLGPGEKYLHAATDTPSVVVHVCSDOENEEEMERMR 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAAGTSTPSPVVHVCSSETEEBE-----T 211

US-09-782-953-22
; Sequence 22, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-22

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFNSP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 ENWDLSDLTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFNSKP 97
QY 62 LSAADARLRHLKTEFLGKEMKLYFAQTLHG---SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINVDLLYAKLGPGEKYLHAATDTPSVVVHVCSDOENEEEMERMR 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAAGTSTPSPVVHVCSSETEEBE-----T 211

US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-23

Query Match          56.8%; Score 595; DB 10; Length 212;
Best Local Similarity 62.6%; Pred. No. 3.8e-53;
```

```
QY 178 KRPKKIIQTRRPE 191
DB 212 KNPKQKIAQTRRPD 225

RESULT 14
US-09-782-953-22
; Sequence 22, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-22

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFNSP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 ENWDLSDLTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFNSKP 97
QY 62 LSAADARLRHLKTEFLGKEMKLYFAQTLHG---SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINVDLLYAKLGPGEKYLHAATDTPSVVVHVCSDOENEEEMERMR 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAAGTSTPSPVVHVCSSETEEBE-----T 211

US-09-782-953-24
; Sequence 24, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-24

Query Match          56.8%; Score 595; DB 10; Length 212;
Best Local Similarity 62.6%; Pred. No. 3.8e-53;
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 136.563 Seconds  
(without alignments)  
934.784 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDQLDLPSTAIACHLDP.....RPKPIQTRPEYTPHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

1: /cgn2\_6/ptodata/1/paa/US06\_COMB.pcp.\*

2: /cgn2\_6/ptodata/1/paa/US07\_COMB.pcp.\*

3: /cgn2\_6/ptodata/1/paa/US08\_COMB.pcp.\*

4: /cgn2\_6/ptodata/1/paa/US09\_COMB.pcp.\*

5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pcp.\*

6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pcp.\*

7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pcp.\*

8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pcp.\*

9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pcp.\*

10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pcp.\*

11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pcp.\*

12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pcp.\*

13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pcp.\*

14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pcp.\*

15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pcp.\*

16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pcp.\*

17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pcp.\*

18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pcp.\*

19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pcp.\*

20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pcp.\*

21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pcp.\*

22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pcp.\*

23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pcp.\*

24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pcp.\*

25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pcp.\*

26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pcp.\*

27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1048	100.0	198	19	US-09-575-580B-4
2	1048	100.0	198	21	US-09-782-953-3
3	1048	100.0	198	21	US-09-782-953-4
4	1002.5	95.7	197	21	US-09-782-953-15
5	1002.5	95.7	197	21	US-09-782-953-16
6	1002.5	95.7	252	25	US-10-104-047-2892

7	920.5	87.8	198	21	US-09-782-953-6	Sequence 6, Appli
8	920.5	87.8	198	21	US-09-782-953-7	Sequence 7, Appli
9	884.5	84.4	197	19	US-09-575-580B-8	Sequence 8, Appli
10	872.5	83.3	197	21	US-09-782-953-12	Sequence 12, Appli
11	872.5	83.3	197	21	US-09-782-953-13	Sequence 13, Appli
12	872.5	83.3	197	26	US-10-247-671-174	Sequence 174, App
13	872.5	83.3	197	27	US-09-323-784-174	Sequence 174, App
14	843.5	80.5	170	19	US-09-575-580B-7	Sequence 7, Appli
15	620.5	59.2	234	24	US-10-030-613-1	Sequence 1, Appli
16	620.5	59.2	234	27	US-09-142-678-1	Sequence 1, Appli
17	620.5	59.2	243	25	US-10-104-047-2216	Sequence 2216, Ap
18	620.5	59.2	255	20	US-09-614-474-2	Sequence 2, Appli
19	613.5	58.5	197	21	US-09-782-953-9	Sequence 9, Appli
20	613.5	58.5	197	21	US-09-782-953-10	Sequence 10, Appli
21	609.5	58.2	192	19	US-09-575-580B-6	Sequence 6, Appli
22	609.5	58.2	197	1	PCT-US02-17382-132	Sequence 132, App
23	606.5	57.9	242	19	US-09-575-580B-5	Sequence 5, Appli
24	596.5	56.9	192	20	US-09-614-474-10	Sequence 10, Appli
25	596.5	56.9	192	21	US-09-782-953-18	Sequence 18, Appli
26	596.5	56.9	192	21	US-09-782-953-19	Sequence 19, Appli
27	596	56.9	241	1	PCT-US02-17382-133	Sequence 133, App
28	596	56.9	241	20	US-09-614-474-11	Sequence 11, Appli
29	596	56.9	241	21	US-09-782-953-21	Sequence 21, Appli
30	596	56.9	242	21	US-09-782-953-22	Sequence 22, Appli
31	595	56.8	212	21	US-09-782-953-24	Sequence 24, Appli
32	595	56.8	212	21	US-09-782-953-25	Sequence 25, Appli
33	583.5	55.7	236	19	US-09-575-580B-24	Sequence 24, Appli
34	497	47.4	142	1	PCT-US00-05918-790	Sequence 790, App
35	497	47.4	142	23	US-09-325-302-790	Sequence 790, App
36	432	41.2	142	21	US-09-758-472-6466	Sequence 6466, Ap
37	432	41.2	142	26	US-10-235-926-6466	Sequence 6466, Ap
38	379.5	36.2	292	20	US-09-614-150-41193	Sequence 41193, A
39	379.5	36.2	292	20	US-09-619-049-1440	Sequence 1440, Ap
40	379.5	36.2	292	20	US-09-167-324-718	Sequence 718, App
41	379.5	36.2	292	27	US-09-171-627-2109	Sequence 2109, App
42	379.5	36.2	292	27	US-09-173-386-680	Sequence 680, App
43	379.5	36.2	292	27	US-09-175-871-764	Sequence 764, App
44	379.5	36.2	292	27	US-09-184-775-692	Sequence 692, App
45	379.5	36.2	292	27	US-09-191-637-40818	Sequence 40818, A

ALIGNMENTS

RESULT 1

US-09-575-580B-4

; Sequence 4, Application US/09575580B

; GENERAL INFORMATION:

; APPLICANT: McKeon, F.

; APPLICANT: Kayako, K.

; APPLICANT: Ryom, S.

; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN

; FILE REFERENCE: HMV-048.01

; CURRENT APPLICATION NUMBER: US/09/575.580B

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 4

; LENGTH: 198

; TYPE: PRT

; ORGANISM: Mus-musculus

US-09-575-580B-4

Query Match 100.0%; Score 1048; DB 19; Length 198;  
Best Local Similarity 100.0%; Pred. No. 7.2e-96;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEEVDQLDLPSTAIACHLDP	RVFDVGLCRKFESLFTYDKDTTFQYKSKFRVIRFESN	60
Db	1	MEEVDQLDLPSTAIACHLDP	RVFDVGLCRKFESLFTYDKDTTFQYKSKFRVIRFESN	60
Qy	61	PLSAADARLRLHKTFLGKEMKLYFAQT	LHIGSSHLAPPNDKQFLISPPASPPVGKQV	120

```

Db      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
Qy      181 KPKIIOTRRPEYTPIHLS 198
      181 KPKIIOTRRPEYTPIHLS 198
Db

```

# RESULT 2

```

US-09-782-953-3
; Sequence 3, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

```

```

Query Match      100.0%; Score 1048; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.2e-96;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
Db      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
Db      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
Qy      181 KPKIIOTRRPEYTPIHLS 198
      181 KPKIIOTRRPEYTPIHLS 198
Db

```

## RESULT 3

```

US-09-782-953-4
; Sequence 4, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus

```

## US-09-782-953-4

```

Query Match      100.0%; Score 1048; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.2e-96;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
Db      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
Qy      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
Db      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
Qy      181 KPKIIOTRRPEYTPIHLS 198
      181 KPKIIOTRRPEYTPIHLS 198
Db      181 KPKIIOTRRPEYTPIHLS 198

```

## RESULT 4

```

US-09-782-953-15
; Sequence 15, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

```

```

Query Match      95.7%; Score 1002.5; DB 21; Length 197;
Best Local Similarity 96.5%; Pred. No. 2.5e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
      1 MEEVDLQDLPSATICHADPRVFDGLCRAKFESLFRYYDKDTTFQYFKSPKRVIRINFSN 60
Db      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
      61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPNPDKOFLISPPASPPVGMKQV 120
Qy      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 180
Db      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 179
      121 EDATPVINVDLLVAISKLGPGKYEELHAATDPTPSVVVHVCSDDNEEBEEMERMRKP 179
Qy      181 KPKIIOTRRPEYTPIHLS 198
      181 KPKIIOTRRPEYTPIHLS 198
Db      180 KPKIIOTRRPEYTPIHLS 197

```

## RESULT 5

```

US-09-782-953-16
; Sequence 16, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus

```

```
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match      95.7%; Score 1002.5; DB 21; Length 197;
Best Local Similarity 96.5%; Pred. No. 2.5e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSATIAHLDPVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MEEVDLQDLPSATIAHLDPVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 179
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 KPILQTRRPEYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 180 KPILQTRRPEYTPHLS 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 6
US-10-104-047-2892
; Sequence 2892, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2892
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2892

Query Match      95.7%; Score 1002.5; DB 25; Length 252;
Best Local Similarity 96.5%; Pred. No. 3.6e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSATIAHLDPVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 56 MEEVDLQDLPSATIAHLDPVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 115
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 116 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 175
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 176 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 234
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 KPILQTRRPEYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 235 KPILQTRRPEYTPHLS 252
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
RESULT 7
US-09-782-953-6
; Sequence 6, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match      87.8%; Score 920.5; DB 21; Length 198;
Best Local Similarity 89.9%; Pred. No. 4e-83;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VDLQDLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDRDFSYNFSSLIACVANDDVSESETRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 EDATPVINDLLYAISKLGPGEKLYELHAATDTPSPVVHVCSDDQNEEBEEMERMRP 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 KPILQTRRPEYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 KPILQTRRPEYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-09-782-953-7
; Sequence 7, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match      87.8%; Score 920.5; DB 21; Length 198;
Best Local Similarity 89.9%; Pred. No. 4e-83;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VDLQDLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDRDFSYNFSSLIACVANDDVSESETRAKFESLRTYDKDTTFOYFKSFKVRINFNS 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```

Db 61 PLASADARLRLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
Oy 121 EATPVPINDDLVAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQENEEBEEEMERMKRP 180
Db 121 EATPVPINDDLVAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQENEEBEEEMERMKRP 180
Oy 181 KPKIIQTRRREYTPHLS 198
Db 181 KPKIIQTRRREYTPHLS 198

```

RESULT 9  
US-09-575-580B-8  
Sequence 8, Application US/09575580B  
GENERAL INFORMATION:  
APPLICANT: MCKEON, F.  
APPLICANT: KAYAKO, K.  
TITLE OF INVENTION: CALCIIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN  
FILE REFERENCE: HMV-048.01  
CURRENT APPLICATION NUMBER: US/09/575,580B  
CURRENT FILING DATE: 2000-05-22  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: Patent In Ver. 2.1  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Citicellus griseus  
US-09-575-580B-8

Query Match 84.4%; Score 884.5; DB 19; Length 197;  
Best Local Similarity 91.5%; Pred. No. 1.6e-79;  
Matches 172; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

```

Oy 11 SATIACHLDPRVFVNDGLCRAPFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLR 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLRQ 70
Oy 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Db 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Oy 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQENEEBEEEMERMKRPKPKIIQTRRP 190
Db 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQEN-EEBEEEMERMKRPKPKIIQTRRP 189
Oy 191 EYTPHLS 198
Db 190 EYTPHLS 197

```

RESULT 10  
US-09-782-953-12  
Sequence 12, Application US/09782953  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: UTSD:674P21  
CURRENT APPLICATION NUMBER: US/09/782,953  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/216,601  
PRIOR FILING DATE: 2000-07-07  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 12  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-782-953-12

```

Query Match 83.3%; Score 872.5; DB 21; Length 197;
Best Local Similarity 89.4%; Pred. No. 2.5e-76;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
Oy 11 SATIACHLDPRVFVNDGLCRAPFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLR 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLRQ 70
Oy 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Db 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Oy 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQENEEBEEEMERMKRPKPKIIQTRRP 190
Db 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQ-KEEBEEMERMRPKPKIIQTRRP 189
Oy 191 EYTPHLS 198
Db 190 EYTPHLS 197

```

RESULT 11  
US-09-782-953-13  
Sequence 13, Application US/09782953  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: UTSD:674P21  
CURRENT APPLICATION NUMBER: US/09/782,953  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/216,601  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 13  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-782-953-13

Query Match 83.3%; Score 872.5; DB 21; Length 197;  
Best Local Similarity 89.4%; Pred. No. 2.5e-76;  
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

```

Oy 11 SATIACHLDPRVFVNDGLCRAPFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLR 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITVDKDTTFQYFKSFKVRINFSNPISADARLRQ 70
Oy 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Db 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINVD 130
Oy 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQENEEBEEEMERMKRPKPKIIQTRRP 190
Db 131 LLYAISKLGPGKEKYLHAATDPTPSVVVHVCSDDQ-KEEBEEMERMRPKPKIIQTRRP 189
Oy 191 EYTPHLS 198
Db 190 EYTPHLS 197

```

RESULT 12  
US-10-247-671-174  
Sequence 174, Application US/10247671  
GENERAL INFORMATION:  
APPLICANT: MIKITA, THOMAS  
APPLICANT: SHIFFMAN, DOV  
APPLICANT: PORTER, GORDON, J.  
APPLICANT: KAISER, MATTHEW R.  
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
US-10-247-671-174

Query Match 59.2%; Score 620.5; DB 24; Length 234;  
Best Local Similarity 62.4%; Pred. No. 4.8e-53;  
Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

```

Oy 5 DIQDPSATIACHDPRVVDGLCRAKFESLPRTYDXTTFQYFKSPKRVIRINFSNPLSA 64
Db 42 DNDLPSNLSLFCACNVHQSVEGGESEKKEFGLFRTYDDCVTFQLFKSPRRVIRINFSNPKSA 101
Oy 65 ADARLRLHKTEFLGKEMKLYFAOTLHIGSS---HLAPPNPDKQPLISPPASPPVGMKOV 120
Db 102 ARARIEHETOFRGKKLKYFAOVOTPETDGDKLHLAPQPAKQPLISPPSSPPVGMQPI 161
Oy 121 EDATPVINYDLLYAISKLGPGKEYELHAATDPTPSVYVAVCESDOENEEEMERMKRP 180
Db 162 NDATPVINYDLLYAVAKLGPGKEYELHAGTESTPSVYVAVCDSDIEEDPK-----TSP 216
Oy 181 KPKIOTRRREYTP 194
Db 217 KPKIOTRRRGLP 230

```

Search completed: December 11, 2002, 11:49:33  
 Job time : 137.563 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 ; Search time 11.6863 Seconds  
(without alignments)  
1118.077 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQDLPSTAIACHLDP.....RPKPIIOTRRPEYPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 233758 seqs, 65991041 residues

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New.\*  
1: /cgn2\_6/prodata1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	999.5	95.4	197	5	US-09-724-676-61352
2	999.5	95.4	197	5	US-09-724-676-61352
3	872.5	83.3	197	5	US-09-724-676-61358
4	872.5	83.3	197	5	US-09-724-676-61359
5	872.5	83.3	197	5	US-09-724-676-61367
6	872.5	83.3	197	5	US-09-724-676-61358
7	872.5	83.3	197	5	US-09-724-676-61359
8	872.5	83.3	197	5	US-09-724-676-61357
9	848	80.9	200	5	US-09-724-676-61356
10	848	80.9	200	5	US-09-724-676-61356
11	847.5	80.9	171	5	US-09-724-676-61354
12	847.5	80.9	171	5	US-09-724-676-61354
13	731	69.8	154	5	US-09-724-676-61353
14	731	69.8	154	5	US-09-724-676-61353
15	620.5	59.2	255	6	US-10-290-438-2
16	607	57.9	205	5	US-09-724-676-61360
17	607	57.9	205	5	US-09-724-676-61360
18	604	57.6	154	5	US-09-724-676-61361
19	604	57.6	154	5	US-09-724-676-61361
20	603.5	57.6	162	5	US-09-724-676-61351
21	603.5	57.6	162	5	US-09-724-676-61351
22	596.5	56.9	192	6	US-10-290-438-10
23	596	56.9	241	6	US-10-290-438-11
24	580	55.3	156	5	US-09-724-676-61355
25	580	55.3	156	5	US-09-724-676-61355
26	579.5	55.3	157	5	US-09-724-676-61357

27 579.5 55.3 157 5 US-09-724-676A-61357 Sequence 61357, A  
28 578.5 55.2 184 5 US-09-724-676-63940 Sequence 63940, A  
29 578.5 55.2 184 5 US-09-724-676A-63940 Sequence 63940, A  
30 439 41.9 111 5 US-09-513-999C-5849 Sequence 5849, A  
31 415.5 39.6 96 5 US-09-724-676-61363 Sequence 61363, A  
32 415.5 39.6 96 5 US-09-724-676-61366 Sequence 61366, A  
33 415.5 39.6 96 5 US-09-724-676A-61363 Sequence 61363, A  
34 415.5 39.6 96 5 US-09-724-676A-61366 Sequence 61366, A  
35 284 27.1 58 6 US-10-203-138A-12463 Sequence 12463, A  
36 271.5 25.9 56 6 US-10-203-138A-11146 Sequence 11146, A  
37 150 14.3 104 5 US-09-724-676-61364 Sequence 61364, A  
38 150 14.3 104 5 US-09-724-676A-61364 Sequence 61364, A  
39 146.5 14.0 61 5 US-09-724-676-61365 Sequence 61365, A  
40 146.5 14.0 61 5 US-09-724-676A-61365 Sequence 61365, A  
41 81 7.7 396 5 US-09-724-676-68208 Sequence 68208, A  
42 81 7.7 396 5 US-09-724-676A-68208 Sequence 68208, A  
43 81 7.7 434 5 US-09-724-676-68213 Sequence 68213, A  
44 81 7.7 434 5 US-09-724-676A-68213 Sequence 68213, A  
45 81 7.7 440 5 US-09-724-676-68200 Sequence 68200, A

## ALIGNMENTS

RESULT 1  
US-09-724-676-61352  
; Sequence 61352, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61352  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-61352

Query Match 95.4%; Score 999.5; DB 5; Length 197;  
Best Local Similarity 96.0%; Pred. No. 2.1e-94;  
Matches 190; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSTAIACHLDPVFDGLCRAPESLRTYDKDTTFQYKSFKEVRINFSN 60  
Db 1 MEEVDLQDLPSTAIACHLDPVFDGLCRAPESLRTYDKDTTFQYKSFKEVRINFSN 60  
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGWKQV 120  
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGWKQV 120  
QY 121 EDATPVINDLYIAIKLGGKGYELHAATDTPSVVVHVCSDDQNEEEEEEMERMRP 180  
Db 121 EDATPVINDLYIAIKLGGKGYELHAATDTPSVVVHVCSDDQNEEEEEEMERMRP 179  
QY 181 KPKIIOTRRPEYPIHLS 198  
Db 180 KPKIIOTRRPEYPIHLS 197

RESULT 2  
US-09-724-676A-61352  
; Sequence 61352, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 61352
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61352

```

```

Query Match          95.4%; Score 999.5; DB 5; Length 197;
Best Local Similarity 96.0%; Pred. No. 2.1e-94;
Matches 190; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 1 MEEVLOLPSTATICHDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MEEVLEDPSTATICHDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 PLASADARLRHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 PPSADARLRHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 EDATFVINDLLYAIKSLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRP 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 EDATFVINDLLYAIKSLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRP 179
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 181 KPRTIQTREPEYPIHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 180 KPRTIQTREPEYPIHLS 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 3
US-09-724-676-61358
; Sequence 61358, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61358
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61358

```

```

Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

```

```

QY 11 SATIACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLR 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 11 SSLIACVANSDFISSETRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLRQ 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 190
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 189
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 191 EYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 190 EYTPHLS 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 4
US-09-724-676-61359
; Sequence 61359, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61359

```

```

Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

```

```

QY 11 SATIACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLR 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 11 SSLIACVANSDFISSETRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLRQ 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 190
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 189
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 191 EYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 190 EYTPHLS 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 5
US-09-724-676-61367
; Sequence 61367, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61367
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61367

```

```

Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

```

```

QY 11 SATIACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLR 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 11 SSLIACVANSDFISSETRAKFESLFRITYDKDTTFQYFKSPKRVINFSNPLSADARLRQ 70
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVIND 130
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 190
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 131 LLYAISKLGPGSEKYLHAATDPTPSVVHVHVCESDQNEEBEEMERMRPKPKTIQTRRP 189
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 191 EYTPHLS 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 190 EYTPHLS 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 6
US-09-724-676A-61358
; Sequence 61358, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.3524 Seconds  
(without alignments)  
513.170 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MFRPFSYNFSLIACVAND.....RPKPKIIQTRPEYTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep:\*

2: /cgn2\_6/prodata/1/iaa/5B-COMB.pep:\*

3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep:\*

4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep:\*

5: /cgn2\_6/prodata/1/iaa/PCUS-COMB.pep:\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835.5	80.1	171	2	US-08-665-040-2
2	84.5	8.1	719	4	US-09-641-741-28
3	82	7.9	284	5	PCT-US94-09752-4
4	82	7.9	1128	4	US-09-641-741-30
5	82	7.9	1128	4	US-09-060-482-8
6	81	7.8	1128	1	US-08-111-939-2
7	81	7.8	2237	1	US-08-455-543A-48
8	81	7.8	2237	2	US-08-223-305C-48
9	81	7.8	2237	4	US-09-268-163-8
10	81	7.8	2339	1	US-08-455-543A-47
11	81	7.8	2339	2	US-08-223-305C-47
12	81	7.8	2339	4	US-09-268-163-6
13	81	7.8	2343	4	US-09-268-163-4
14	80	7.7	370	2	US-08-231-342-6
15	79.5	7.6	434	2	US-08-815-718-3
16	79.5	7.6	559	4	US-08-884-072-6
17	79.5	7.6	559	4	US-09-213-168-6
18	78.5	7.5	201	2	US-08-716-317-1
19	78.5	7.5	547	4	US-09-877-730-26
20	78.5	7.5	624	4	US-09-877-730-24
21	78.5	7.5	628	4	US-09-877-730-30
22	78.5	7.5	712	4	US-09-877-730-22
23	78.5	7.5	793	4	US-09-877-730-28
24	78.5	7.5	826	4	US-09-877-730-16
25	78.5	7.5	904	4	US-09-877-730-6
26	78.5	7.5	907	4	US-09-877-730-20
27	78.5	7.5	985	4	US-09-877-730-10

28	78.5	7.5	991	4	US-09-877-730-12
29	78.5	7.5	1059	4	US-09-877-730-2
30	78.5	7.5	1072	4	US-09-877-730-18
31	78.5	7.5	1150	4	US-09-877-730-8
32	78.5	7.5	1706	2	US-08-459-568-2
33	78.5	7.5	1706	2	US-08-399-411-2
34	78.5	7.5	1706	3	US-08-516-859A-2
35	78.5	7.5	1706	4	US-09-586-472-2
36	78.5	7.5	1706	4	US-09-528-706-2
37	77	7.4	522	6	RE34606-6
38	77	7.4	1912	4	US-08-913-832A-2
39	77	7.4	1912	4	US-09-249-181A-2
40	77	7.4	2337	3	US-08-713-118-2
41	77	7.4	2337	4	US-09-452-090-2
42	75	7.2	362	2	US-08-985-090-5
43	75	7.2	362	3	US-09-165-543-32
44	75	7.2	1158	4	US-09-060-482-2
45	74	7.1	445	3	US-09-165-543-5

## ALIGNMENTS

## RESULT 1

US-08-665-040-2

; Sequence 2, Application US/08665040

; Patent No. 5869318

; GENERAL INFORMATION:

; APPLICANT: ESTIVILL PALLEJA, XAVIER

; APPLICANT: FUENTES, JUAN JOSE

; APPLICANT: PRITCHARD, MELANIE

; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE

; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,

; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,

; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY

; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LADAS & PARRY

; STREET: 26 WEST 61ST STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10023

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORD PERFECT 5.1 FOR DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/665,040

; FILING DATE: JUNE 7, 1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: ES P9501140

; FILING DATE: JUNE 7, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: JANET I. CORD

; REGISTRATION NUMBER: 33,778

; REFERENCE/DOCKET NUMBER: U010815-9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 708-1800

; TELEFAX: (212) 246-8959

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 171 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-665-040-2

Query Match 80.1%; Score 835.5; DB 2; Length 171;  
 Best Local Similarity 94.7%; Pred. No. 4,4e-87;  
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Qy 30 AKFESLFRTYDQTTQYFKSKRVYRINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 89  
 Db 4 AKFESLFRTYDQTTQYFKSKRVYRINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 63  
 Qy 90 HIGSSHLAPNDKQFLSPSPVGVKQVADATPVINYDLVAISKLGPEKXELHAA 149  
 Db 64 HIGSSHLAPNDKQFLSPSPVGVKQVADATPVINYDLVAISKLGPEKXELHAA 123  
 Qy 150 TDPTSVVHVCESDQENEEEMERMKRPKRIIOTRRPEYTPHLS 198  
 Db 124 TDTTSVVHVCESDQENEEEMERMKRPKRIIOTRRPEYTPHLS 171

RESULT 2  
 US-09-641-741-28  
 Sequence 28, Application US/09641741  
 Patent No. 6420155

GENERAL INFORMATION:  
 APPLICANT: Kerry E. Quinn  
 TITLE OF INVENTION: Aortic Carboxypeptidase-like Proteins and Nucleic Acids  
 TITLE OF INVENTION: encoding Same  
 FILE REFERENCE: 15966-581  
 CURRENT APPLICATION NUMBER: US/09/641,741  
 CURRENT FILING DATE: 2000-08-18  
 PRIOR APPLICATION NUMBER: 60/159,613  
 PRIOR FILING DATE: 1999-10-14  
 PRIOR APPLICATION NUMBER: 60/175,534  
 PRIOR FILING DATE: 2000-01-12  
 PRIOR APPLICATION NUMBER: 60/224,086  
 PRIOR FILING DATE: 2000-08-09  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 28  
 LENGTH: 719  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-641-741-28

Query Match 8.1%; Score 84.5; DB 4; Length 719;  
 Best Local Similarity 21.8%; Pred. No. 0.71;  
 Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

Qy 54 VRINFSNPLSADARLRLHKEFLGKEMKLYFAQTLHIGSSHLAPNDKQFLSP 110  
 Db 585 LRVDPSRPMTPQORRMQRLQYLRMRBQML--RLNLTAGPATSPSP--ALMP 637  
 Qy 111 ASP-----PVGMK--QVEDATPVINYDLVAISKLGPEKXELHAA 153  
 Db 638 PSTPTATLTPMVLPTTTGMESEETETVTVT-----EFETRYGTD-- 601  
 Qy 154 PSVVHVCESDQENEEEMERME 175  
 Db 682 ----LVEVEIEDEEBEEMD 699

RESULT 3  
 PCT-US94-09752-4  
 Sequence 4, Application PC/TUS9409752  
 GENERAL INFORMATION:  
 APPLICANT: David S. Strayer and Avinash Chander  
 TITLE OF INVENTION: Compositions and Methods for  
 TITLE OF INVENTION: Targeting Cells and Modulating Pulmonary Surfactant Secretion  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Jane Massey Licata, Esq.  
 STREET: 210 Lake Drive East, Suite 201  
 CITY: Cherry Hill  
 STATE: NJ

COUNTRY: USA  
 ZIP: 08002  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb  
 MEDIUM TYPE: STORAGE  
 OPERATING SYSTEM: IBM PS/2  
 SOFTWARE: WORDPERFECT 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/09752  
 FILING DATE: Herewith  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/176,218  
 FILING DATE: December 30, 1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/114,951  
 FILING DATE: August 31, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jane Massey Licata  
 REGISTRATION NUMBER: 32,257  
 REFERENCE/DOCKET NUMBER: JEFF-0042  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (609) 779-8488  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 284  
 TYPE: amino acid  
 TOPOLOGY: linear  
 PCT-US94-09752-4

Query Match 7.9%; Score 82; DB 5; Length 284;  
 Best Local Similarity 24.5%; Pred. No. 0.35;  
 Matches 39; Conservative 23; Mismatches 57; Indels 40; Gaps 7;

Qy 34 SLFRYDQTTQYFKSKRVYRINFSNPLSADARLRLHKEFLGKEMKLYFAQTLHIGS 93  
 Db 10 SLFRSGRQTLFA-----SVMRYAD--LYAASFVN--LTVPPSYLFR 50  
 Qy 94 SHLAPNP-----DKQFLSPSPAS--PPVGMKQVEDATPVINYDLVAISKLGPE 142  
 Db 51 AHVLPHESTVHTHTVDINEMESPLATRNRTSVDFKD---TDYRHOGLTSSISIKKPN 106  
 Qy 143 KYELHAATDPTPSVVHVCESDQENEEEMERMKRPK 181  
 Db 107 LPPL-----APQIETHCHDEDDDEEEBEECCGKPK 139

RESULT 4  
 US-09-641-741-30  
 Sequence 30, Application US/09641741  
 Patent No. 6420155

GENERAL INFORMATION:  
 APPLICANT: Kerry E. Quinn  
 TITLE OF INVENTION: Aortic Carboxypeptidase-like Proteins and Nucleic Acids  
 TITLE OF INVENTION: encoding Same  
 FILE REFERENCE: 15966-581  
 CURRENT APPLICATION NUMBER: US/09/641,741  
 CURRENT FILING DATE: 2000-08-18  
 PRIOR APPLICATION NUMBER: 60/159,613  
 PRIOR FILING DATE: 1999-10-14  
 PRIOR APPLICATION NUMBER: 60/175,534  
 PRIOR FILING DATE: 2000-01-12  
 PRIOR APPLICATION NUMBER: 60/224,086  
 PRIOR FILING DATE: 2000-08-09  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 30  
 LENGTH: 1128  
 TYPE: PRT  
 ORGANISM: Mus musculus

US-09-641-741-30

Query Match 7.9%; Score 82; DB 4; Length 1128;  
Best Local Similarity 20.8%; Pred. No. 2.6;  
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRLHKTFF---LGKMKLYFAOTLHIGS 93  
DB 977 NWKIREILAMNGNRPILGVDPSRPMTPQORRMQORRLQYRLRMREQMRL---RRLNSTA 1033  
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136  
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081  
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175  
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

## RESULT 5

US-09-060-482-8

; Sequence 8, Application US/09060482

; Patent No. 6468766

; GENERAL INFORMATION:

; APPLICANT: Lee, Wu-En

; APPLICANT: Layne, Matthew D.

; APPLICANT: Yet, Shaw-Fang

; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE

; FILE REFERENCE: 05433/036001

; CURRENT APPLICATION NUMBER: US/09/060,482

; EARLIER FILING DATE: 1998-04-15

; EARLIER APPLICATION NUMBER: US 081818,009

; EARLIER FILING DATE: 1997-03-14

; EARLIER APPLICATION NUMBER: US 60\013,439

; EARLIER FILING DATE: 1996-03-15

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 1128

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-060-482-8

Query Match 7.9%; Score 82; DB 4; Length 1128;  
Best Local Similarity 20.8%; Pred. No. 2.6;  
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRLHKTFF---LGKMKLYFAOTLHIGS 93  
DB 977 NWKIREILAMNGNRPILGVDPSRPMTPQORRMQORRLQYRLRMREQMRL---RRLNSTA 1033  
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136  
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081  
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175  
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

## RESULT 6

US-08-111-939-2

; Sequence 2, Application US/08111939

; Patent No. 5460951

; GENERAL INFORMATION:

; APPLICANT: Kawai, Shinji

; APPLICANT: Takeshita, Sunao

; APPLICANT: Okazaki, Makoto

; APPLICANT: Amann, Egon

; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like

; TITLE OF INVENTION: Protein and Process for its Production

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER: IBM PC compatible  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/111,939  
FILING DATE: 26-AUG-1993  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 324033/92

FILING DATE: 03-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 230029/92

FILING DATE: 28-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: Forman, David S.

REGISTRATION NUMBER: 33,694

REFERENCE/DOCKET NUMBER: 02481.1321-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1128 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-111-939-2

Query Match 7.8%; Score 81; DB 1; Length 1128;

Best Local Similarity 20.8%; Pred. No. 3.4;

Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRLHKTFF---LGKMKLYFAOTLHIGS 93  
DB 977 NWKIREILAMNGNRPILGVDPSRPMTPQORRMQORRLQYRLRMREQMRL---RRLNSTA 1033  
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136  
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081  
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175  
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

## RESULT 7

US-08-455-543A-48

; Sequence 48, Application US/08455543A

; Patent No. 5792846

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

; APPLICANT: Brenner, Robert

; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller &amp; McClain

STREET: 1660 Union Street

CITY: San Diego

STATE: California

COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,543A  
FILING DATE: May 31, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/223,305  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-52517  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-455-543A-48

Query Match 7.8%; Score 81; DB 1; Length 2237;  
Best Local Similarity 21.5%; Pred. No. 9.4; Indels 75; Gaps 9;  
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFSSESETRAKFESLFTYDXTTFQYFKSPKRVINFSNPLISAAD 66  
DB 697 NYLLNVFLAIADNLNANQELTRDEEEMBAANQKLAOKAKVAEV-----SPMSAAN 751  
QY 67 -----ARLRH---KTEFLGKEN---KLYFAOTLHIGSSHLA 97  
DB 752 ISIAHQNSAKARSWEORASQRLRQNLNASCALYSEMDPEERLRFATTRL-----805  
QY 98 PRPDQOFLSP-----ASPPVGMKQVEDATPVINYLVAISKLGPEXYELHA 148  
DB 806 --RPDMKTHLDRLVVELGDAAGPVGAKAPEAAE-----APEGVDPRRHHRR 855  
QY 149 ATPPTSVVHVCSDE-----NEEEEMEMKRPKP 182  
DB 856 DKDKTPAA-----GQDDRAEAPKAESGEGPQAEERPRP 888

RESULT 8

US-08-223-305C-48  
Sequence 48, Application US/08223305C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/223,305C  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-223-305C-48

Query Match 7.8%; Score 81; DB 2; Length 2237;  
Best Local Similarity 21.5%; Pred. No. 9.4; Indels 75; Gaps 9;  
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFSSESETRAKFESLFTYDXTTFQYFKSPKRVINFSNPLISAAD 66  
DB 697 NYLLNVFLAIADNLNANQELTRDEEEMBAANQKLAOKAKVAEV-----SPMSAAN 751

QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97  
Db 752 ISIAARQONSAKARSVWEORASQLRNLNLRASCEALYSEMDPEERLRFATTHL-----805  
QY 98 PPNPDQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148  
Db 806 --RPDMKTHLDRPLVVLGRDGGARGPVGGKARPEAAE-----APEGVDPFRHRHR 855  
QY 149 ATDPTPSVVHVCSDOE-----NEEEEMERMKRKP 182  
Db 856 DKDKTPAA-----GQDRAEAPKASGEPGAREPRP 888

RESULT 9  
US-09-268-163-8  
; Sequence 8, Application US/09268163B  
; Patent No. 6353091  
; GENERAL INFORMATION:  
; APPLICANT: Lipscombe, Diane  
; APPLICANT: Schorge, Stephanie  
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF  
; FILE REFERENCE: B1055/7000  
; CURRENT APPLICATION NUMBER: US/09/268,163B  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: US 60/077,901  
; EARLIER FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2237  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-268-163-8

Query Match 7.8%; Score 81; DB 4; Length 2237;  
Best Local Similarity 21.5%; Pred. No. 9.4;  
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFESETRAKFESLFRITYDKDTTFQYKFSKRVIRFNSNPLSAAD 66  
Db 697 NYTLNVLAIADVNLANAQELTKDEEEMEAANQKALQKAEVAEV-----SPMSAAN 751  
QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97  
Db 752 ISIAARQONSAKARSVWEORASQLRNLNLRASCEALYSEMDPEERLRFATTHL-----805  
QY 98 PPNPDQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148  
Db 806 --RPDMKTHLDRPLVVLGRDGGARGPVGGKARPEAAE-----APEGVDPFRHRHR 855  
QY 149 ATDPTPSVVHVCSDOE-----NEEEEMERMKRKP 182  
Db 856 DKDKTPAA-----GQDRAEAPKASGEPGAREPRP 888

RESULT 10  
US-08-455-543A-47  
; Sequence 47, Application US/08455543A  
; Patent No. 5792846  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: Feldman, Daniel  
; APPLICANT: McCue, Ann  
; APPLICANT: Brenner, Robert  
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street

CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,543A  
FILING DATE: May 31, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/223,305  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-52517  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2339 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-455-543A-47

Query Match 7.8%; Score 81; DB 1; Length 2339;  
Best Local Similarity 21.5%; Pred. No. 10;  
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFESETRAKFESLFRITYDKDTTFQYKFSKRVIRFNSNPLSAAD 66  
Db 697 NYTLNVLAIADVNLANAQELTKDEEEMEAANQKALQKAEVAEV-----SPMSAAN 751  
QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97  
Db 752 ISIAARQONSAKARSVWEORASQLRNLNLRASCEALYSEMDPEERLRFATTHL-----805  
QY 98 PPNPDQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148  
Db 806 --RPDMKTHLDRPLVVLGRDGGARGPVGGKARPEAAE-----APEGVDPFRHRHR 855  
QY 149 ATDPTPSVVHVCSDOE-----NEEEEMERMKRKP 182  
Db 856 DKDKTPAA-----GQDRAEAPKASGEPGAREPRP 888

RESULT 11  
 US-08-223-305C-47  
 Sequence 47, Application US/08223305C  
 Patent No. 5851824  
 GENERAL INFORMATION:  
 APPLICANT: Harpold, Michael  
 APPLICANT: Ellis, Steven  
 APPLICANT: Williams, Mark  
 APPLICANT: Feldman, Daniel  
 APPLICANT: McCue, Ann  
 APPLICANT: Brenner, Robert  
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
 TITLE OF INVENTION: METHODS  
 NUMBER OF SEQUENCES: 57  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Brown, Martin, Haller & McClain  
 STREET: 1660 Union Street  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92101-22926  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/223,305C  
 FILING DATE: April 4, 1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/868,354  
 FILING DATE: April 10, 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/745,206  
 FILING DATE: 15-AUG-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/620,250  
 FILING DATE: 30-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/482,384  
 FILING DATE: 20-FEB-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/603,751  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US89/01408  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/176,899  
 FILING DATE: 04-APR-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 52516 (P519739)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 238-0999  
 TELEFAX: (619) 238-0062  
 INFORMATION FOR SEQ ID NO: 47:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2339 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 US-08-223-305C-47

Query Match 7.8%; Score 81; DB 2; Length 2339;  
 Best Local Similarity 21.5%; Pred. No. 10;  
 Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;  
 7 SYNFSLLACVANDVSESESETRAKFESLFRTYDKDTTFQYFKSKRVIRINFSNPLSAAD 66

Db 697 NYTLNVFLAIVDNLNAOELTKDEEEMERANOKLQAKAEVAEY-----SPMSAN 751  
 QY 67 -----ARLRH-----KTEFLGKEM-----KLYPEOTLHIGSSHLA 97  
 Db 752 ISIAARQONSAAKARSWEQASQLRLQNLRASCEALYSEMDPEERLRFATRRHL----- 805  
 QY 98 PPNPDQKFLISPP-----ASPPVGWKQVEDATPVINTDLYAISKLGPGKYTELHA 148  
 Db 806 --RPDMKTHLDRPLVVELGRDARGPVGKARPEAAE-----ABEGVDPRRHRRHR 855  
 QY 149 ATDPTPSVYVHVCSDOE-----NEEESEEMERMKRPP 182  
 Db 856 DKDKTPAA-----GDQDRAEAPKAESEGPARERRPP 888

RESULT 12  
 US-09-268-163-6  
 Sequence 6, Application US/09268163B  
 Patent No. 6353091  
 GENERAL INFORMATION:  
 APPLICANT: Lipescombe, Diane  
 TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF  
 FILE REFERENCE: B1055/7000  
 CURRENT APPLICATION NUMBER: US/09/268,163B  
 CURRENT FILING DATE: 1999-03-12  
 EARLIER APPLICATION NUMBER: US 60/077,901  
 EARLIER FILING DATE: 1998-03-13  
 NUMBER OF SEQ ID NOS: 28  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 6  
 LENGTH: 2339  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-268-163-6

Query Match 7.8%; Score 81; DB 4; Length 2339;  
 Best Local Similarity 21.5%; Pred. No. 10;  
 Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLACVANDVSESESETRAKFESLFRTYDKDTTFQYFKSKRVIRINFSNPLSAAD 66  
 Db 697 NYTLNVFLAIVDNLNAOELTKDEEEMERANOKLQAKAEVAEY-----SPMSAN 751  
 QY 67 -----ARLRH-----KTEFLGKEM-----KLYPEOTLHIGSSHLA 97  
 Db 752 ISIAARQONSAAKARSWEQASQLRLQNLRASCEALYSEMDPEERLRFATRRHL----- 805  
 QY 98 PPNPDQKFLISPP-----ASPPVGWKQVEDATPVINTDLYAISKLGPGKYTELHA 148  
 Db 806 --RPDMKTHLDRPLVVELGRDARGPVGKARPEAAE-----ABEGVDPRRHRRHR 855  
 QY 149 ATDPTPSVYVHVCSDOE-----NEEESEEMERMKRPP 182  
 Db 856 DKDKTPAA-----GDQDRAEAPKAESEGPARERRPP 888

RESULT 13  
 US-08-268-163-4  
 Sequence 4, Application US/09268163B  
 Patent No. 6353091  
 GENERAL INFORMATION:  
 APPLICANT: Lipescombe, Diane  
 TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF  
 FILE REFERENCE: B1055/7000  
 CURRENT APPLICATION NUMBER: US/09/268,163B  
 CURRENT FILING DATE: 1999-03-12  
 EARLIER APPLICATION NUMBER: US 60/077,901  
 EARLIER FILING DATE: 1998-03-13  
 NUMBER OF SEQ ID NOS: 28  
 SOFTWARE: FastSeq for Windows Version 3.0



QY 13 LIACVANDVPESESETRAKFESLFRTYDKDFTTFQYFKSPKRVRFNSNPLSAADA-RLRL 71

Db 10 ILACIALASASEGAFKASDQ--REMTPERLFQHLH-----EVGYAPPSLPQTRRLRV 61  
QY 72 HKTEFLGKEMKLYFAOTLHIGSSHLAPNPDQFLISPPASPPVGWKQVEDATPVINYD- 130  
Db 62 D-----HSTVSLHDPPLFEQREYQPPSP-----ED-IPVEEDM 96  
QY 131 ---LYAISKLGPEKYEELAAATDPTPSVVVAVCESDQENEBEEMERMKRPPKXILOT 187  
Db 97 PTFLLNPVNDKAGPAVFOEAIPLQKEQPPQVHI-----EKEIDPPAQFOEELVQK 147  
QY 188 RREXT 193  
Db 148 EVKPHI 153

Search completed: December 11, 2002, 11:39:15  
Job time : 14.3524 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 ; Search time 11.6863 Seconds  
(without alignments)  
1118.077 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFDSYNFSLIACVAND.....RPKPKLIQTRRPEYTPHLS 198

Scoring table: BLOSUM62

Searched: 233758 seqs, 65991041 residues

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New.\*  
1: /cgn2\_6/prodata1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	970.5	93.0	197	5	US-09-724-676-61358
2	970.5	93.0	197	5	US-09-724-676-61359
3	970.5	93.0	197	5	US-09-724-676-61367
4	970.5	93.0	197	5	US-09-724-676A-61358
5	970.5	93.0	197	5	US-09-724-676A-61359
6	970.5	93.0	197	5	US-09-724-676A-61367
7	874	83.8	197	5	US-09-724-676-61352
8	874	83.8	197	5	US-09-724-676A-61352
9	847.5	81.3	171	5	US-09-724-676-61354
10	847.5	81.3	171	5	US-09-724-676A-61354
11	847.5	81.3	200	5	US-09-724-676-61356
12	847.5	81.3	200	5	US-09-724-676A-61356
13	705	67.6	205	5	US-09-724-676-61360
14	705	67.6	205	5	US-09-724-676A-61360
15	702	67.3	154	5	US-09-724-676-61361
16	702	67.3	154	5	US-09-724-676A-61361
17	701.5	67.3	162	5	US-09-724-676-61351
18	701.5	67.3	162	5	US-09-724-676A-61351
19	623	59.7	255	6	US-10-290-438-2
20	615.5	59.0	192	6	US-10-290-438-10
21	605.5	58.1	154	5	US-09-724-676-61353
22	605.5	58.1	154	5	US-09-724-676A-61353
23	597.5	57.3	184	5	US-09-724-676-63940
24	597.5	57.3	184	5	US-09-724-676A-63940
25	592.5	56.8	241	6	US-10-290-438-11
26	580	55.6	156	5	US-09-724-676-61355

27	580	55.6	156	5	US-09-724-676A-61355	Sequence 61355, A
28	579	55.5	157	5	US-09-724-676-61357	Sequence 61357, A
29	579	55.5	157	5	US-09-724-676A-61357	Sequence 61357, A
30	537	51.5	111	5	US-09-513-999C-5849	Sequence 5849, Ap
31	415.5	39.8	96	5	US-09-724-676-61363	Sequence 61363, A
32	415.5	39.8	96	5	US-09-724-676-61366	Sequence 61366, A
33	415.5	39.8	96	5	US-09-724-676A-61363	Sequence 61363, A
34	415.5	39.8	96	5	US-09-724-676A-61366	Sequence 61366, A
35	284	27.2	58	6	US-10-203-138A-12463	Sequence 12463, A
36	271.5	26.0	56	6	US-10-203-138A-11146	Sequence 11146, A
37	150	14.4	104	5	US-09-724-676-61364	Sequence 61364, A
38	150	14.4	104	5	US-09-724-676A-61364	Sequence 61364, A
39	146.5	14.0	61	5	US-09-724-676-61365	Sequence 61365, A
40	146.5	14.0	61	5	US-09-724-676A-61365	Sequence 61365, A
41	127.5	12.2	80	5	US-09-724-676-61362	Sequence 61362, A
42	127.5	12.2	80	5	US-09-724-676A-61362	Sequence 61362, A
43	81	7.8	396	5	US-09-724-676-68208	Sequence 68208, A
44	81	7.8	396	5	US-09-724-676A-68208	Sequence 68208, A
45	81	7.8	434	5	US-09-724-676-68213	Sequence 68213, A

ALIGNMENTS

RESULT 1  
US-09-724-676-61358  
; Sequence 61358, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724.676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61358  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-61358

Query Match 93.0%; Score 970.5; DB 5; Length 197;  
Best Local Similarity 93.4%; Pred. No. 1.6e-89;  
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy	1	MDRFDYSNFSLLIACVANDVSESETRAKFESLFRYDKDITFOYKSPKRVINFSN	60
Db	1	MHFRNFYSFSLIACVANSDFSESETRAKFESLFRYDKDITFOYKSPKRVINFSN	60
Qy	61	PLSAADARLRHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGVKQV	120
Db	61	PLSAADARLRHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGVKQV	120
Qy	121	EDATPVINYDLYLAISKLGPKYKELHAATDTPSVVHVHVCESDOENEEEMERMRP	180
Db	121	EDATPVINYDLYLAISKLGPKYKELHAATDTPSVVHVHVCESDOE-KEEEMERMRP	179
Qy	181	KPKLIQTRRPEYTPHLS	198
Db	180	KPKLIQTRRPEYTPHLS	197

RESULT 2  
US-09-724-676-61359  
; Sequence 61359, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724.676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 61359
;
; LENGTH: 197
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
US-09-724-676-61359

```

Query Match	93.0%;	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%;	Pred. No. 1,6e-69;		
Matches 185; Conservative	7;	Mismatches 5;	Indels 1;	Gaps 1;

[illegible]

RESULT 3  
US-09-724-676-61367

TITLE OF INVENTION: Variants of alternative splicing  
 FILE REFERENCE: 129181.4 Compugen  
 CURRENT APPLICATION NUMBER: US/09/724,676  
 CURRENT FILING DATE: 2000-11-28  
 NUMBER OF SEQ ID NOS: 97222  
 SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 61367
;
; LENGTH: 197
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-09-724-676-61367

```

Query Match	93.0%	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%	Pred. No. 1.6e-89;		
Matches 185; Conservative	7;	Mismatches 5;	Indels 1;	Gaps 1

```

Oy 1 MDPRFSNPSLLICVANDDVFSSESEFRAPFESI,FRYYDKLTFFQYFKSKRIRINFSN 60
Db 1 MHEKPNFNSFSSLLICVANSDDI,FSSESEFRAPFESI,FRYYDKDITFFQYFKSKRIRINFSN 60
Oy 61 PLSAADARLR,LRHKTEFLGKEMKLYPAQTLHGSSHLAPBNEDKQFLISPPASPVGKQY 120
Db 61 PLSAADARLR,LRHKTEFLGKEMKLYPAQTLHGSSHLAPBNEDKQFLISPPASPVGKQY 120
Oy 121 EDATPVINDLLIYAI,SKUGPGEKYE,LAATDP,TPS,VVVHV,VESNQEN,EBE,EBE,ERNAKRP 180
Db 121 EDATPVINDLLIYAI,SKUGPGEKYE,LAATDP,TPS,VVVHV,VESNQDQ-,KEEBE,EBE,MRMRP 179
Oy 181 KPKIIQTRRPEYTPIHLS 198
Db 180 KPKIIQTRRPEYTPIHLS 197

```

RESULT 4  
US-09-724-676A-61358

```

: Sequence 6158, Application US/09724676A
:
: GENERAL INFORMATION:
:   APPLICANT: CompuGen LTD
:   TITLE OF INVENTION: Variants of alternative splicing
:   FILE REFERENCE: 129181.1 CompuGen
:   CURRENT APPLICATION NUMBER: US/09/774,676A

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
;

```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-61358

Query Match	93.0%	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%	Pred. No. 1.6e-89;		
Matches 185; Conservative	7;	Mismatches 5;	Indels 1;	Gaps 1;

```

QY      1 MDPDFSVNPSLLACVANDVFESESTRKPSBLPFTYKDTFOYKSKFKYRINFSN 60
Db      1 MAFRFNFSVSSLLACVANSDDISESETRAKFBSLFRITDKDITFOYKSKFKYRINFSN 60
QY      61 PLSAADARLRHAKTEFLGKEMKLYFAQTLLIGSSHLAPPNPDQKFLISPPASPVGVKQY 120
Db      61 PPSAADALQHKTEFLGKEMKLYFAQTLLIGSSHLAPPNPDQKFLISPPASPVGVKQY 120
QY      121 EDATPVINVDLLVAISKGPGEKTELHAIDPTPSVVVHVHCESDQNEEBEEMERMKRP 180
Db      121 EDATPVINVDLLVAISKGPGEKTELHAIDPTPSVVVHVHCESDQNEEBEEMERMKRP 179
QY      161 KPKIQTRRPEYTPILHS 198
Db      160 KPKIQTRRPEYTPILHS 197

```

RESULT 5  
US-09-724-676A-61359

```

1 TITLE OF INVENTION: Variants of alternative splicing
2
3 FILE REFERENCE: 129181.4 Compugen
4
5 CURRENT APPLICATION NUMBER: US/09/724,676A
6
7 CURRENT FILING DATE: 2000-11-28
8
9 NUMBER OF SEQ ID NOS: 97222
10
11 SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61359

```

Query Match	93.0%	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%	Pred. No. 1.6e-89;		
Matches 185;	Conservative	7;	Mismatches 5;	Indels 1;
				Gaps 1;

```
Qy      1 MOFREYNFSLACVANDVDFESESRAFEELPTTYOKDTFOYKFKKKVRINFN 60
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1 MHFRFNISFSLLACVANSIDISESETRAKFESLFTTYDKDITFOYKFKPKVRINFN 60
        |||||
Qy      61 PL$ADARLRILHKEPFLECKEMKLYFAOTLHGSSHLAPPNPDQFLISPPASPVGNKY 120
        |||||
Db      61 P$SADARLRQLHKEPFLCKEMKLYFAOTLHGSSHLAPNPDPQFLISPPASPVGNKY 120
        |||||
Qy      121 EDATPVINVDLLIYAISKLGPEKXELHAADPTPSVVVHVHCSDQENBEEBEEMERMKRP 180
        |||||
Db      121 EDATPVINVDLLIYAISKLGPEKXELHAADPTTPSVVHVHCSDQE-KEBEEEMERM RP 179
        |||||
Qy      181 KPKIIQTRRREPTPIHLS 198
        |||||
Db      180 KPKIIQTRRREPTPIHLS 197
```

RESULT 6  
US-09-724-676A-61367

```

RESULT 6
US-09-72A-676A-61367
; Sequence 61367, Application US/0972A676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
;

```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-61358

Query Match	93.0%	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%	Pred. No. 1.6e-89;		
Matches 185; Conservative	7;	Mismatches 5;	Indels 1;	Gaps 1;

```

QY      1 MDPDFSVNPSLLACVANDVFESESTRKPSBLPFTYKDTFOYKSKFKYRINFN 60
Db      1 MHRFRNFSSLLACVANSDDISESETRAKFBSLFRITDKDITFOYKSKFKYRINFN 60
QY      61 PLSAADARLRHAKTEFLGKEMKLYFAQTLLIGSSHLAPPNPDQKFLISPPASPVGVKQY 120
Db      61 PPSAADALQHKTEFLGKEMKLYFAQTLLIGSSHLAPPNPDQKFLISPPASPVGVKQY 120
QY      121 EDATPVINVDLLVAISKSGPEKTELHAADPTPSVVVHVHCESDQNEEBEEMERKRP 180
Db      121 EDATPVINVDLLVAISKSGPEKTELHAADPTPSVVVHVHCESDQNEEBEEMERKRP 179
QY      181 KPKIQTRRPEYTPILHS 198
Db      180 KPKIQTRRPEYTPILHS 197

```

RESULT 5  
US-09-724-676A-61359

```

1 TITLE OF INVENTION: Variants of alternative splicing
2
3 FILE REFERENCE: 129181.4 Compugen
4
5 CURRENT APPLICATION NUMBER: US/09/724,676A
6
7 CURRENT FILING DATE: 2000-11-28
8
9 NUMBER OF SEQ ID NOS: 97222
10
11 SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61359

```

Query Match	93.0%	Score 970.5;	DB 5;	Length 197;
Best Local Similarity	93.4%	Pred. No. 1.6e-89;		
Matches 185;	Conservative	7;	Mismatches 5;	Indels 1;
				Gaps 1;

```
Qy      1 MOFREYNFSLACVANDVDFESESRAFEELPTTYOKDTFOYKFKKKVRINFN 60
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1 MHFRFNYSFSLLACVANSIDISESETRAKFESLFTTYDKDITFOYKFKPKVRINFN 60
        |||||
Qy      61 PL$ADARLRHKTEFLGCKEMKLYFAOTLHGSSHLAPPNPDQKFLISPPASPVGNKY 120
        |||||
Db      61 P$SADARLRHKTEFLGCKEMKLYFAOTLHGSSHLAPPNPDQKFLISPPASPVGNKY 120
        |||||
Qy      121 EDATPVIIVDLIYAISKLGPEKYTELHAADTPTSVVVHVHCESDOENBEEBEEMERMKRP 180
        |||||
Db      121 EDATPVIIVDLIYAISKLGPEKYTELHAADTPTSVVVHVHCESDOE-KEBEEEMERM RP 179
        |||||
Qy      181 KPKTIQTRRREPTYPIHLS 198
        |||||
Db      180 KPKTIQTRRREPTYPIHLS 197
```

RESULT 6  
US-09-724-676A-61367

```

RESULT 6
US-09-72A-676A-61367
; Sequence 61367, Application US/0972A676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD

```

; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 61367  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61367

Query Match 93.0%; Score 970.5; DB 5; Length 197;  
Best Local Similarity 93.4%; Pred. No. 1.6e-89;  
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;  
QY 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
DB 1 MHFNFYSFSSLIACVANDSISESETRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
QY 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
DB 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
QY 181 KPKIOTRRPEYTPHLS 198  
DB 180 KPKIOTRRPEYTPHLS 197

RESULT 7  
US-09-724-676-61352  
; Sequence 61352, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 61352  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-61352

Query Match 83.8%; Score 874; DB 5; Length 197;  
Best Local Similarity 86.4%; Pred. No. 7.1e-80;  
Matches 171; Conservative 7; Mismatches 16; Indels 4; Gaps 2;  
QY 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
DB 4 VLEEDLP---SATIACHLDPRVFDGLCRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
QY 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
DB 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
QY 181 KPKIOTRRPEYTPHLS 198  
DB 180 KPKIOTRRPEYTPHLS 197

RESULT 8  
US-09-724-676A-61352

; Sequence 61352, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 61352  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61352

Query Match 83.8%; Score 874; DB 5; Length 197;  
Best Local Similarity 86.4%; Pred. No. 7.1e-80;  
Matches 171; Conservative 7; Mismatches 16; Indels 4; Gaps 2;  
QY 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
DB 4 VLEEDLP---SATIACHLDPRVFDGLCRAKFESLFTYDKDTTFQYFKSFKVRINFSN 60  
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGKQV 120  
QY 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
DB 121 EDATPVINYDLLVAISKLGEGEYELHAATDTPSVVVHVHVCSDQNEEEEEEMRMKRP 180  
QY 181 KPKIOTRRPEYTPHLS 198  
DB 180 KPKIOTRRPEYTPHLS 197

RESULT 9  
US-09-724-676-61354  
; Sequence 61354, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 61354  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-61354

Query Match 81.3%; Score 847.5; DB 5; Length 171;  
Best Local Similarity 95.9%; Pred. No. 2.6e-77;  
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
QY 30 AKFESLFTYDKDTTFQYFKSFKVRINFSNPLSADARLRLHKTFLGKEMKLYFAQTL 89  
DB 4 AKFESLFTYDKDTTFQYFKSFKVRINFSNPLSADARLRLHKTFLGKEMKLYFAQTL 63  
QY 90 HIGSSHLAPPNDKQFLISPPASPPVGKQVEDATPVINYDLLVAISKLGEGEYELHAA 149  
DB 64 HIGSSHLAPPNDKQFLISPPASPPVGKQVEDATPVINYDLLVAISKLGEGEYELHAA 123  
QY 150 TDPTPSVVHVHVCSDQNEEEEEEMRMKRPKPKIOTRRPEYTPHLS 198  
DB 124 TDPTPSVVHVHVCSDQNEEEEEEMRMKRPKPKIOTRRPEYTPHLS 171

RESULT 10  
US-09-724-676A-61354  
; Sequence 61354, Application US/09724676A

```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61354
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61354

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 171;
Best Local Similarity 95.9%; Pred. No. 2,6e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 4 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 63
Qy 90 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 149
Db 64 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 123
Qy 150 TDPPTSVVVHVCSDOENEEMERMRKRPKIIQTRREPTPIHLS 198
Db 124 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIIQTRREPTPIHLS 171

```

```

RESULT 11
US-09-724-676-61356
; Sequence 61356, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61356
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61356

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 200;
Best Local Similarity 95.9%; Pred. No. 3,2e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 33 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 92
Qy 90 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 149
Db 93 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 152
Qy 150 TDPPTSVVVHVCSDOENEEMERMRKRPKIIQTRREPTPIHLS 198
Db 153 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIIQTRREPTPIHLS 200

```

```

RESULT 12
US-09-724-676A-61356
; Sequence 61356, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61356
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61356

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 200;
Best Local Similarity 95.9%; Pred. No. 3,2e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 33 AKFESLFRYYDKOTTFQYFKSFKRVRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 92
Qy 90 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 149
Db 93 HIGSSLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 152
Qy 150 TDPPTSVVVHVCSDOENEEMERMRKRPKIIQTRREPTPIHLS 198
Db 153 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIIQTRREPTPIHLS 200

```

```

RESULT 13
US-09-724-676-61360
; Sequence 61360, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61360
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61360

```

```

Query Match      67.6%; Score 705; DB 5; Length 205;
Best Local Similarity 93.0%; Pred. No. 5,7e-63;
Matches 133; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 1 MDRFDSYNSFLIACVANDDVSESETRAKFESLFRYYDKOTTFQYFKSFKRVRINFSN 60
Db 1 MAFRPNYSFSSLIACVANSDFSESETRAKFESLFRYYDKOTTFQYFKSFKRVRINFSN 60
Qy 61 PLSADARLRLHKTFFLGKEMKLYFAQTLHIGSSLAPPNPDKQFLISPPASPPVGMKQV 120
Db 61 PLSADARLRLHKTFFLGKEMKLYFAQTLHIGSSLAPPNPDKQFLISPPASPPVGMKQV 120
Qy 121 EDPATPVINYDLVLAISKLGPEK 143
Db 121 EDPATPVINYDLVLAISKLGPEQ 143

```

```

RESULT 14
US-09-724-676A-61360
; Sequence 61360, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61360
; LENGTH: 205

```



THIS PAGE BLANK (UBPTO)



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:09:21 ; Search time 52.7555 Seconds  
(without alignments)  
2732.361 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFSYNFSSLIACVAND.....RPKPKIIQRRPEYPIHLS 198

Scoring table: BLSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODES=frame+ p2n.model -DB=xlh  
-Q=/cgn2\_1/USPTO.spool/US09782953/runat\_11122002\_114431\_17431/app\_query.fasta\_1.1173  
-DB=Pending Patents NA New -QFWT=fastcap -SUFFIX=inn -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 @CNC 1.170 @runat\_11122002\_114431\_17431 -NCPU=6 -ICPU=3  
-NO XUPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:

1: /cgn2\_6/prodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/prodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	970.5	93.0	1931	5	US-09-724-676-12748 Sequence 12748, A
2	970.5	93.0	1931	5	US-09-724-676A-12748 Sequence 12748, A
3	970.5	93.0	2050	5	US-09-724-676-12756 Sequence 12756, A
4	970.5	93.0	2050	5	US-09-724-676A-12756 Sequence 12756, A
5	970.5	93.0	2297	5	US-09-724-676-12747 Sequence 12747, A
6	970.5	93.0	2297	5	US-09-724-676A-12747 Sequence 12747, A
7	957.5	91.8	2348	6	US-10-240-965-60 Sequence 60, Appl
8	874	83.8	1893	5	US-09-724-676-12741 Sequence 12741, A
9	874	83.8	1893	5	US-09-724-676A-12741 Sequence 12741, A
10	850.5	81.5	1876	5	US-09-724-676-12745 Sequence 12745, A
11	850.5	81.5	1876	5	US-09-724-676A-12745 Sequence 12745, A

12	847.5	81.3	1839	5	US-09-724-676-12743 Sequence 12743, A
13	847.5	81.3	1839	5	US-09-724-676A-12743 Sequence 12743, A
14	705	67.6	911	5	US-09-724-676-12749 Sequence 12749, A
15	705	67.6	911	5	US-09-724-676A-12749 Sequence 12749, A
16	702	67.3	1094	5	US-09-724-676-12750 Sequence 12750, A
17	702	67.3	1094	5	US-09-724-676A-12750 Sequence 12750, A
18	702	67.3	1213	5	US-09-724-676-12740 Sequence 12740, A
19	702	67.3	1213	5	US-09-724-676A-12740 Sequence 12740, A
20	623	59.7	3159	6	US-10-290-438-1 Sequence 1, Appl
21	605.5	58.1	1056	5	US-09-724-676-12742 Sequence 12742, A
22	605.5	58.1	1056	5	US-09-724-676A-12742 Sequence 12742, A
23	597.5	57.3	2411	5	US-09-724-676-15329 Sequence 15329, A
24	597.5	57.3	2411	5	US-09-724-676A-15329 Sequence 15329, A
25	582	55.8	1039	5	US-09-724-676-12746 Sequence 12746, A
26	582	55.8	1039	5	US-09-724-676A-12746 Sequence 12746, A
27	579	55.5	1002	5	US-09-724-676-12744 Sequence 12744, A
28	579	55.5	1002	5	US-09-724-676A-12744 Sequence 12744, A
29	537	51.5	442	5	US-09-513-999C-1772 Sequence 1772, Ap
30	536.5	51.4	615	6	US-10-290-438-8 Sequence 8, Appl
31	528.5	50.7	1021	6	US-10-290-438-3 Sequence 3, Appl
32	415.5	39.8	1577	5	US-09-724-676-12755 Sequence 12755, A
33	415.5	39.8	1577	5	US-09-724-676A-12755 Sequence 12755, A
34	415.5	38.8	1943	5	US-09-724-676-12752 Sequence 12752, A
35	415.5	38.8	1943	5	US-09-724-676A-12752 Sequence 12752, A
36	285.5	27.4	446	6	US-10-203-138A-2101 Sequence 2101, Ap
37	266.5	25.6	486	6	US-10-203-138A-827 Sequence 827, App
38	198.5	19.0	531	6	US-10-290-438-6 Sequence 6, Appl
39	187.5	18.0	123	6	US-10-203-138A-5980 Sequence 5980, Ap
40	170	16.3	465	5	US-09-620-607B-1320 Sequence 1320, Ap
41	150	14.4	557	5	US-09-724-676-12753 Sequence 12753, A
42	150	14.4	557	5	US-09-724-676A-12753 Sequence 12753, A
43	147	14.1	740	5	US-09-724-676-12754 Sequence 12754, A
44	147	14.1	740	5	US-09-724-676A-12754 Sequence 12754, A
45	140	13.4	85	6	US-10-203-138A-7216 Sequence 7216, Ap

#### ALIGNMENTS

RESULT 1  
US-09-724-676-12748  
; Sequence 12748, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: CompuGen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 CompuGen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12748  
; LENGTH: 1931  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12748

Alignment Scores:  
Pred. No.: 4.31e-100 Length: 1931  
Score: 970.50 Matches: 185  
Percent Similarity: 96.97% Conservative: 7  
Best Local Similarity: 93.43% Mismatches: 5  
Query Match: 93.05% Indels: 1

DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676-12748 (1-1931)

QY 1 MetAapPheArGAspPheSerTyRAsnPheSerSerLeuIleAlaCyVaValAlaAsnAsp 20  
 DB 63 ATGCATTTTAGAACTTTAACTTACATTAAGCTCCCTGATTCCTGCTGGCAACAGT 122

QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrAsp 40  
 DB 123 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGAGTCCCTCTTAGAGCATGAC 182

QY 41 LysAapThrThrPheGlnTyRPhelySerPheLysValArgIleAsnPheSerAsn 60  
 DB 183 AAGGACATACCTTCAGATTTTAAAGACTTCAAAAGAGTCAGAAATTAACCTTCAGCAAC 242

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluLysGlu 80  
 DB 243 CCTCTCTCCGAGCAGATGCCAGGCTCCAGCTGATAGACTGAGTTTCTGGAAAGGA 302

QY 81 MetLysLeuTyRPhelAglnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
 DB 303 ATGAACTTATATTTGCTCAGACCTTACATAGAAAGCTCACACTGGCTCCGCCAAAT 362

QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120  
 DB 363 CCAGACAAAGCTTTCATCTCCCTCCGCTCCGCCAGTGGGATGGAAACAAAGTG 422

QY 121 GluAapAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140  
 DB 423 GAAGAGCCAGCCCACTCATTAACATATGATCTTTATAGCCATCTCCAAAGCTGGGCA 482

QY 141 GlyLysLysTyRGlulLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
 DB 483 GGGGAAAGTATGAAATTCAGCAGCAGCTGACACCACTCCAGCGTGGTGCATGTA 542

QY 161 CygGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180  
 DB 543 TGTGAAGTGCATCAAGAG--AAGGAGGAAGAAAGAAATGGAATGAGAGACCT 599

QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198  
 DB 600 AAGCCAAAATTTATCCAGACCAAGAGCGGAGTACCGCCATCCACTCAGC 653

RESULT 2  
 US-09-724-676A-12748  
 ; Sequence 12748, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724, 676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12748  
 ; LENGTH: 1931  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (18)..(18)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (18)..(18)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (23)..(23)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676A-12748

Alignment Scores:

Pred. No.:	4,31e-100	Length:	1931
Score:	970.50	Matches:	185
Percent Similarity:	96.97%	Conservative:	7
Best Local Similarity:	93.43%	Mismatches:	5
Query Match:	93.05%	Indels:	1
DB:	5	Gaps:	1

US-09-782-953-7 (1-198) x US-09-724-676A-12748 (1-1931)

QY 1 MetAapPheArGAspPheSerTyRAsnPheSerSerLeuIleAlaCyVaValAlaAsnAsp 20  
 DB 63 ATGCATTTTAGAACTTTAACTTACATTAAGCTCCCTGATTCCTGCTGGCAACAGT 122

QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrAsp 40  
 DB 123 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGAGTCCCTCTTAGAGCATGAC 182

QY 41 LysAapThrThrPheGlnTyRPhelySerPheLysValArgIleAsnPheSerAsn 60  
 DB 183 AAGGACATACCTTCAGATTTTAAAGACTTCAAAAGAGTCAGAAATTAACCTTCAGCAAC 242

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluLysGlu 80  
 DB 243 CCTCTCTCCGAGCAGATGCCAGGCTCCAGCTGATAGACTGAGTTTCTGGAAAGGA 302

QY 81 MetLysLeuTyRPhelAglnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
 DB 303 ATGAACTTATATTTGCTCAGACCTTACATAGAAAGCTCACACTGGCTCCGCCAAAT 362

QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120  
 DB 363 CCAGACAAAGCTTTCATCTCCCTCCGCTCCGCCAGTGGGATGGAAACAAAGTG 422

QY 121 GluAapAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140  
 DB 423 GAAGTGGCAGCCCACTCATTAACATATGATCTTTATAGCCATCTCCAAAGCTGGGCA 482

QY 141 GlyLysLysTyRGlulLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
 DB 483 GGGGAAAGTATGAAATTCAGCAGCAGCTGACACCACTCCAGCGTGGTGCATGTA 542

QY 161 CygGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180  
 DB 543 TGTGAAGTGCATCAAGAG--AAGGAGGAAGAAAGAAATGGAATGAGAGACCT 599

QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198  
 DB 600 AAGCCAAAATTTATCCAGACCAAGAGCGGAGTACCGCCATCCACTCAGC 653

RESULT 3  
 US-09-724-676-12756  
 ; Sequence 12756, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724, 676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12756  
 ; LENGTH: 2050  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (7)..(7)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (47)..(47)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676A-12748

US-09-724-676-12756

**Alignment Scores:**

Pred. No.:	4.68-100	Length:	2050
Score:	970.50	Matches:	185
Percent Similarity:	96.97%	Conservative:	7
Best Local Similarity:	93.43%	Mismatches:	5
Query Match:	93.05%	Indels:	1
DB:	5	Gaps:	1

US-09-782-953-7 (1-198) x US-09-724-676-12756 (1-2050)

Qy	1	MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp	20
Db	182	ATGCATTTTGAAGAACTTAACATACAGCTTTTAGCTCCCTGATTCCTGTGTGGCAACAGT	241
Qy	21	AspValPheSerGluSerGluThrArgAlaValPheGluSerLeuPheArgThrTyrAsp	40
Db	242	GATATCTTCAGCGNAAGTGAACCCAGGGCCAAATTTAGTGCCCTCTTTAGGACGTATGAC	301
Qy	41	LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn	60
Db	302	AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGAAATAAATTCAGCAAC	361
Qy	61	ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu	80
Db	362	CCCTTCTCCGACGAGATGCGAGCTCCAGCTCCATAGAAGACTGAGTTTCTGGAAAGGAA	421
Qy	81	MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn	100
Db	422	ATGAAGTTATATTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCGCCAAAT	481
Qy	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal	120
Db	482	CCAGACAAGCAGTTCTCGATCTCCCTCCCGCTCTCCGCCAGTGGGATGGAAACAAGTG	541
Qy	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140
Db	542	GAAGATCGACCCCGAGTCATAAACTATGATCTCTTATGCCATCTCCAAAGCTGGGGCCA	601
Qy	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal	160
Db	602	GGGGAAAGTATGATTTGCACGGACGACATGACACCACTCCCAAGCTGGTGGTCCATGTA	661
Qy	161	CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro	180
Db	662	TGTGAGAGTCAACAAG---AAGGAGGAAGAAGAGAAATGGAAAGATGAGGAGACCT	718
Qy	181	LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer	198
Db	719	AAGCCAAAATTTACAGACGAGGCGGAGTACACGCCATCCACCTCAGC	772

## RESULT 4

US-09-724-676A-12756  
; Sequence 12756, Application US/09724676A

```

; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 12756
;
; LENGTH: 2050
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n i
;
; FEATURE:
;
; NAME/KEY: misc_feature

```

US-09-724-676A-12756

Alignment Scores:	
Pred. No.:	4.68e-100
Score:	970.50
Percent Similarity:	66.97%
Best Local Similarity:	93.43%
Query Match:	93.05%
DB:	5
Length:	2050
Matches:	185
Conservative:	7
Mismatches:	5
Indels:	1
Gaps:	1

US-09-782-953-7 (1-198) x US-09-724-676A-12756 (1-2050)

Qy	1	MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp	20
Db	182	ATGCATTTTAAAGAACTTAACTACAGCTTTTAGCTCCCTGATTCCTGTCGGCAACAGT	241
Qy	21	AspValPheSerGluSerGluThrArgAlaLySpheGluSerLeuPheArgThrTyrAsp	40
Db	242	GATATCTTTCAGGAAAGTGAACACAGGCGCAATTTGAGTCCCTCTTTAGGACGTATGAC	301
Qy	41	LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn	60
Db	302	AAGGACATCACCTTCTCAGTATTTTAAGAGCTTCAACAGAGCTCAGAAATNACTTCACGCAAC	361
Qy	61	ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu	80
Db	362	CCCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAAAGACTGAGTTTCTGGGAAAGGAA	421
Qy	81	MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPheAsn	100
Db	422	ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACTGGCTCGCCCAAT	481
Qy	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal	120
Db	482	CCAGACAAGCAGTTCTCGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAGTG	541
Qy	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140
Db	542	GAAGATGCCGACCCGAGTCATAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGSCCA	601
Qy	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal	160
Db	602	GGGGAAAGATATGAATTGCACGGCAGCGACTGCACCACTCCCAAGCTGGTGGTCCATGTA	661
Qy	161	CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro	180
Db	662	TGTGAGAGTGATCAACAG---AAGGAGGAAGAGAGAGAAATGGAAAGAATGAGGAGACCT	718
Qy	181	LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer	198
Db	719	AAGCCAAAATTTATCAGACCGAGGCGGAGTACACGCCGATCCACCTCAGC	772

## RESULT 5

US-09-724-676-12747  
; Sequence 12747, Application US/09724676

```

; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```
; SEQ ID NO 12747
;
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n i
```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (18)..(18)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12747

```

```

Alignment Scores:
Pred. No.: 5,476-100 Length: 2297
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7
Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1
DB: Gaps: 1
US-09-782-953-7 (1-198) x US-09-724-676-12747 (1-2297)

```

```

QY 1 MetAapPheArGAspPheSerTyraenPheSerSerLeuIleAlaCyValAlaAsnAap 20
DB 63 ATGCATTTTAGAAACTTTAACTTACCTTTAGCTCCCTGATTGCTGCTGGCAAAACAGT 122
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40
DB 123 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGAGTCCCTCTTGAAGCGTATGAC 182
QY 41 LysAapThrThrPheGluThrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
DB 183 AAGGACATCACTTTAGATTTTAAAGCTTCAACAGAGTCAAGATTAACCTTCAGCAAC 242
QY 61 ProlSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluLysGlu 80
DB 243 CCTCTTCGCGAGCAATGCGAGCTCCAGCTGCATTAAGCTGAGTTCTGGGAAAGAA 302
QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
DB 303 ATGAACTTATATTGCTCTCAGACCTTACATAGAGAACTGACACCTGCTCCGCCAAAT 362
QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
DB 363 CCGAGCAACAGATTGATGATCTCCCTCCGCTTCCGCAAGTGGAGTGAAGAAACAAGTG 422
QY 121 GluAapAlaThrProValIleAsnTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
DB 423 GAAAGATGCCAGCCCACTCATTAACCTATGATCTTTATATGCCATCTCCAAGCTGGGCCA 482
QY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAapProThrProSerValValHisVal 160
DB 483 GGGGAAAGATATGAATTGCAACGACGACAGTCAACACTCCCACTGGTGGTCCATGTA 542
QY 161 CygGluSerAapGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
DB 543 TGTGAAGAGTATCAAGAG--AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198
DB 600 AAGCCAAAATTTATCCAGACCGAGAGCCGGAGTACAGCCGATCCACCTCAGC 653

```

```

RESULT 6
US-09-724-676A-12747
Sequence 12747, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12747
LENGTH: 2297

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11)..(11)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (18)..(18)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12747

```

```

Alignment Scores:
Pred. No.: 5,476-100 Length: 2297
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7
Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1
DB: Gaps: 1
US-09-782-953-7 (1-198) x US-09-724-676A-12747 (1-2297)

```

```

QY 1 MetAapPheArGAspPheSerTyraenPheSerSerLeuIleAlaCyValAlaAsnAap 20
DB 63 ATGCATTTTAGAAACTTTAACTTACCTTTAGCTCCCTGATTGCTGCTGGCAAAACAGT 122
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40
DB 123 GATATCTTCAGCGAAAGTAAACAGGCGCAATTGAGTCCCTCTTGAAGCGTATGAC 182
QY 41 LysAapThrThrPheGluThrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
DB 183 AAGGACATCACTTTAGATTTTAAAGCTTCAACAGAGTCAAGATTAACCTTCAGCAAC 242
QY 61 ProlSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluLysGlu 80
DB 243 CCTCTTCGCGAGCAATGCGAGCTCCAGCTGCATTAAGCTGAGTTCTGGGAAAGAA 302
QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
DB 303 ATGAACTTATATTGCTCTCAGACCTTACATAGAGAACTGACACCTGCTCCGCCAAAT 362
QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
DB 363 CCGAGCAACAGATTGATGATCTCCCTCCGCTTCCGCAAGTGGAGTGAAGAAACAAGTG 422
QY 121 GluAapAlaThrProValIleAsnTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
DB 423 GAAAGATGCCAGCCCACTCATTAACCTATGATCTTTATATGCCATCTCCAAGCTGGGCCA 482
QY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAapProThrProSerValValHisVal 160
DB 483 GGGGAAAGATATGAATTGCAACGACGACAGTCAACACTCCCACTGGTGGTCCATGTA 542
QY 161 CygGluSerAapGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
DB 543 TGTGAAGAGTATCAAGAG--AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198
DB 600 AAGCCAAAATTTATCCAGACCGAGAGCCGGAGTACAGCCGATCCACCTCAGC 653

```

```

RESULT 7
US-10-240-965-60
Sequence 60, Application US/10240965
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIFFMAN, DOV
APPLICANT: SOMOGYI, Roland

```

APPLICANT: LAWN, Richard M.  
APPLICANT: SEILHAMER, Jeffrey J.  
APPLICANT: PORTER, Gordon J.  
APPLICANT: MIKITA, Thomas  
APPLICANT: TAI, Julie  
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
CURRENT APPLICATION NUMBER: US/10/240,965  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: 60/195,106  
PRIOR FILING DATE: 2000-04-05  
NUMBER OF SEQ ID NOS: 276  
SOFTWARE: PERL Program  
SEQ ID NO 60  
LENGTH: 2348  
TYPE: DNA  
ORGANISM: Homo sapiens  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: 042176.5  
US-10-240-965-60

Alignment Scores:  
Pred. No.: 1,65e-98 Length: 2348  
Score: 957.50 Matches: 185  
Percent Similarity: 96.48% Conservative: 7  
Best Local Similarity: 92.96% Mismatches: 5  
Query Match: 91.80% Indels: 2  
DB: 6 Gaps: 1

US-09-782-953-7 (1-198) x US-10-240-965-60 (1-2348)

```
Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20
Db 151 ATGCATTTTAGAACTTTAACTACAGTTTAGCTCCCTGATTCCTGTGTGGCAACAGT 210
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 211 GATATCTTCAGCGAAGTGAACCCAGGGCCAAATTTGAGTCCCTCTTTAGGAGGTATGAC 270
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsn 60
Db 271 AAGGACATCACCCTTCAGTATTTAAGAGCTTCAACAGCTCAGATTAACCTTCAGCAAC 330
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 331 CCCTTCTCCGACGAGATGCCAGCTCCAGCTCCAGTAAAGACTCAGTTCTCGGAAAGGAA 390
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysSerSerHisLeuAlaProProAs 100
Db 391 ATGAGTTATATTTGCTCAGACCTTACACATAGGAAGCTTCACACTGGGCTCCGCAAA 450
Qy 100 nProAspLysGlnPheLeuLeuSerProAlaSerProProValGlyTyrLysGlnVa 120
Db 451 TCCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTCAGTGGATGGAAACAGT 510
Qy 120 lGluAspAlaThrProValLeuSerAspLeuLeuTyrAlaLysSerLysLeuGlyPr 140
Db 511 GGAAGATGCGACCCAGTATATAAATATGATCTCTTATATGCTCTCAAGCTGGGGCC 570
Qy 140 oGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVa 160
Db 571 AGGGGAAAAGTATGAATTGACACGAGCTGACACCTCCAGGCTGGTGGTCCATCT 630
Qy 160 lCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPr 180
Db 631 ATGTGAGTGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGAC 687
Qy 180 oLysProLysLeuLeuThrArgArgProGluTyrThrProLysLeuSer 198
Db 688 TAAGCAAAAATTTATCCAGACCGAGGCGGAGTACAGCCGCTCCACTCAGC 742
```

RESULT 8

US-09-724-676-12741  
Sequence 12741, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12741  
LENGTH: 1893  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-724-676-12741

Alignment Scores:  
Pred. No.: 3.21e-89 Length: 1893  
Score: 874.00 Matches: 171  
Percent Similarity: 89.90% Conservative: 7  
Best Local Similarity: 86.36% Mismatches: 16  
Query Match: 83.80% Indels: 4  
DB: 5 Gaps: 2

US-09-782-953-7 (1-198) x US-09-724-676-12741 (1-1893)

```
Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20
Db 34 GTGGACCTCGAGGACCTGCCCC-----ACGGCCACCATCGCTGTCCACCTGGACCG 84
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 144
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsn 60
Db 145 AAGGACATCACCCTTTTCAGTATTTAAGAGCTTCAACAGCTCAGAAATTAACCTTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysSerSerHisLeuAlaProProAsn 100
Db 265 ATGAGTTATATTTGCTCAGACCTTACACATAGGAAGCTTCACACTGGCTCCGCAAAAT 324
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 325 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTGGGATGGAAACAAGTG 384
Qy 121 GluAspAlaThrProValLeuSerAspLeuLeuTyrAlaLysSerLysLeuGlyPro 140
Db 395 GAAGATGCGACCCAGTATATAAATATGATCTCTTATATGCTCTCAAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAAAGTATGATTTCCAGCGAGCTCAGACCTCCAGCTCCAGCTGGTGGTCCATGTA 504
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGTGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACT 561
Qy 181 LysProLysLeuLeuGlnThrArgArgProGluTyrThrProLysLeuSer 198
Db 562 AAGCAAAAATTTCCAGACCGAGGCGGAGTACAGCGGCTCCACTCAGC 615
```

RESULT 9

US-09-724-676A-12741  
Sequence 12741, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen

LENGUIN: 10/0

Alignment Scores:  
Pred No : 1.42e-86  
length: 1876

Alignment scores:	
Pred. No.:	1.42e-86
Length:	1876

Score: 850.50 Matches: 165  
Percent Similarity: 93.48% Conservativity: 7  
Best Local Similarity: 89.67% Mismatches: 10  
Query Match: 81.54% Indels: 2  
DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676A-12745 (1-1876)

QY 16 CysValAlaAsnAspValPheSerGluSerGlu-ThrArgAlaLysPheGluSerLe 35  
Db 52 TGCATTCCCAATCGCACTATTGGAAGAAAGCAATGTTGGGCCAAATTTGAGTCCCT 111  
QY 35 upheArgThrTyraAspLysAspThrPheGlnTyPheLysSerPheLysArgValAr 55  
Db 112 CTTTAGACGATGACAAAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAAGAGTCAG 171  
QY 55 gileAsnPheserAsnProLysSerAlaAlaAspAlaArgLeuGluHisLysThrGl 75  
Db 172 AATAAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGA 231  
QY 75 upheLeuGluLysGluMetLysLeuTyPheAlaGlnThrLeuHisLysGlySerHi 95  
Db 232 GTTTCCTGGGAAGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACA 291  
QY 95 sLeuAlaProAsnProAspLysGlnPheLeuLysSerProAlaSerProProVa 115  
Db 292 CTGGCTCCGCCAATCCAGACAGCAAGTTCCTGATCTCCCTCCCGCCCTCTCCGCGAGT 351  
QY 115 lGlyTrpLysGlnValGluAspAlaThrProValLysAsnTyraAspLeuTyraAla 135  
Db 352 GGGATGGAACAAAGTGAAGATGCGACCCAGTCATATAAATATGATCTTATATGCCAT 411  
QY 135 eSerLysLeuGluProGlyLysTyraGluLeuHisAlaAlaThrAspProThrProse 155  
Db 412 CTCACAGCTGGGCCAGGGAAGATGAATTCGACGAGGACTGACACCACTCCCGAG 471  
QY 155 rValValHisValCysGluSerAspGlnGluAsnGluGluGluGluMetGl 175  
Db 472 CTGGTGGTCCATGATGTGAGAGTATCAAGAG---AAGGAGGAAGAGGAAATGA 528  
QY 175 uArgMetLysArgProLysProLysLeuGlnThrArgArgProGluTyraThrProIl 195  
Db 529 AAGATGAGGAGACCTTAAGCAAAATTTATCCAGACGAGGAGCGGAGTACAGCGCGAT 588  
QY 195 eHisLeuSer 198  
Db 589 CCACCTCAGC 598

RESULT 12  
US-09-724-676-12743  
; Sequence 12743, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 12743  
; LENGTH: 1839  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-12743

Alignment Scores:  
Pred. No.: 3e-86 Length: 1839  
Score: 847.50 Matches: 162  
Percent Similarity: 97.63% Conservativity: 3  
Best Local Similarity: 95.86% Mismatches: 3  
Query Match: 81.26% Indels: 1  
DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676-12743 (1-1839)  
QY 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyraPheLys 49  
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAG 117  
QY 50 SerPheLysArgValArgLysLeuHisLysSerAsnProLysSerAlaAlaAspAlaArgLeu 69  
Db 118 AGCTTCAACAGGATCAGATAAATTTACAGAACCCCTTCTCCGACGACATGCCAGGCTC 177  
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyraPheAlaGlnThrLeu 89  
Db 178 CAGCTGCATTAAGCTGAGTTCTTGGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTA 237  
QY 90 HisLysGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109  
Db 238 CACATAGGAGCTCACCTGCTCCGCCAATCCAGACAGCAGTTTCTGATCTCCCT 297  
QY 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValLysAsnTyra 129  
Db 298 CCGGCTCTCCGCCAGTGGGATGGAACAAAGTGAAGATGCGACCCAGTCAATAACTAT 357  
QY 130 AspLeuLeuTyraAlaLysSerLysLeuGlyProGlyGluLysTyraGluLeuHisAla 149  
Db 358 GATCTCTTATATGCTATCTCCAGCTGGGCCAGGGAAGATGATGAATGACGACGCG 417  
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
Db 418 ACTGACACCACTCCAGCTGGTGTCTCATGTATGTGAGATGATCAAGAG---AAGGAG 474  
QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysLeuGlnThrArgArg 189  
Db 475 GAAGAAGAGGAATGGAAGATGAGGACCTTAAGCAAAATTTATCCAGACCAAGGAGG 534  
QY 190 ProGluTyraThrProLysHisLeuSer 198  
Db 535 CCGGAGTACACCGCATCCACCTCAGC 561

RESULT 13  
US-09-724-676A-12743  
; Sequence 12743, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 12743  
; LENGTH: 1839  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676A-12743

Alignment Scores:  
Pred. No.: 3e-86 Length: 1839  
Score: 847.50 Matches: 162  
Percent Similarity: 97.63% Conservativity: 3  
Best Local Similarity: 95.86% Mismatches: 3  
Query Match: 81.26% Indels: 1  
DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676A-12743 (1-1839)  
QY 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyraPheLys 49  
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAG 117  
QY 50 SerPheLysArgValArgLysLeuHisLysSerAsnProLysSerAlaAlaAspAlaArgLeu 69  
Db 118 AGCTTCAACAGGATCAGATAAATTTACAGAACCCCTTCTCCGACGACATGCCAGGCTC 177

Qy 70 ArgLeuHisLeuThrGluPheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeu 89  
Db 178 CAGCTGCACAAAGACTGAGTTCTCGGAGAAAGAAATGAAGTTATATTCTCAGACCTTA 237  
Qy 90 HisIleGlySerSerHisLeuAlaProProAlaProAspLysGlnPheLeuIleSerPro 109  
Db 238 CACATGAGGAAGCTCACACCTGGCTCCGCCAAATTCAGAACACAGTTCTTGATCTCCCT 297  
Qy 110 ProAlaSerProProValGlyTyrPheGluValGluAspAlaThrProValIleAsnTyr 129  
Db 298 CCGGCTCTCCGCGCATGGATGGAAACAACTGGAAGATGCCACCCCATTAACCTTA 357  
Qy 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149  
Db 358 GATCTCTTATATGATCCATCTCCAAGCTGGGCCAGGGGAAAGATGAATTCACACCGACG 417  
Qy 150 ThrAspProThrProSerValValValHisValCysGluSerAspGlnLysGlnGlu 169  
Db 418 ACTGACACCACTCCACGCTGGTGTCCATGTATGTGAGAGATGATCAGAG--AAGGAG 474  
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189  
Db 475 GAGAGAGAGAAATGAGAAAGAAATGAGAGAGACTTACGCCAAATTTATCCAGACCGAGAG 534  
Qy 190 ProGluTyrThrProIleHisLeuSer 198  
Db 535 CCGAGTACACGCCGATCCACCTCAGC 561

## RESULT 14

US-09-724-676-12749  
Sequence 12749, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12749  
LENGTH: 911  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (11)..(11)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (18)..(18)  
OTHER INFORMATION: n is a,c,g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (23)..(23)  
OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12749

## Alignment Scores:

Pred. No.: 1,35e-70 Length: 911  
Score: 705.00 Matches: 133  
Percent Similarity: 97.20% Conservative: 6  
Best Local Similarity: 93.01% Mismatches: 4  
Query Match: 67.59% Indels: 0  
DB: 5 Gaps: 0

US-09-782-953-7 (1-198) x US-09-724-676-12749 (1-911)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20  
Db 63 ATGCAATTTTGAACCTTAACACTAGCTTACCTCCCTGATGCTGTGGCAACAGT 122  
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
Db 123 GATATCTTCAGCGAAGATGAAACCAAGGCCCAATTGAGTCCCTCTTTAGACGATATAC 182

Db 123 GATATCTTCAGCGAAGATGAAACCAAGGCCCAATTGAGTCCCTCTTTAGACGATATAC 182  
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
Db 183 AAGACATACCTCTTCAGCTATTTTAAAGAGCTTCAAGAGATGAGAAATTAATCTCAGAAC 242  
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 243 CCTCTTCGCGACGATGCGCAGCTCCAGCTGCATTAAGCTGATTTCTGGGAAAGAA 302  
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100  
Db 303 ATGAGTATATTTTGTCTCAGACCTTAACATAGAGAGCTCACCTGCTCGCCAAAT 362  
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPheGlnVal 120  
Db 363 CCAAGCAAGCAAGTTTGTGATCTCCCTCCGCTCCGCGACGAGGAGTGGAAACAAAGTG 422  
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 423 GAAGATGCGACCCAGATCATTAACATGATCTTATATGCAATCTCAAGCTGGGCCA 482  
Qy 141 GlyGlyLys 143  
Db 483 GGTGACGAG 491

## RESULT 15

US-09-724-676A-12749  
Sequence 12749, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12749  
LENGTH: 911  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (11)..(11)  
OTHER INFORMATION: n is a,c,g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (18)..(18)  
OTHER INFORMATION: n is a,c,g, or t  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (23)..(23)  
OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12749

## Alignment Scores:

Pred. No.: 1,35e-70 Length: 911  
Score: 705.00 Matches: 133  
Percent Similarity: 97.20% Conservative: 6  
Best Local Similarity: 93.01% Mismatches: 4  
Query Match: 67.59% Indels: 0  
DB: 5 Gaps: 0

US-09-782-953-7 (1-198) x US-09-724-676A-12749 (1-911)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20  
Db 63 ATGCAATTTTGAACCTTAACACTAGCTTACCTCCCTGATGCTGTGGCAACAGT 122  
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
Db 123 GATATCTTCAGCGAAGATGAAACCAAGGCCCAATTGAGTCCCTCTTTAGACGATATAC 182



```
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
    |||||
Db 183 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAAAACGAGTTCAGATATAAACTTCAGCAAC 242
    |||||
Qy 61 ProLeuSerAlaAlaAspAlaAtgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
    |||||
Db 243 CCTTCTCCGAGCAGATGCGAGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAA 302
    |||||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
    |||||
Db 303 ATGAAGTTATATTTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 362
    |||||
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
    |||||
Db 363 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAGTG 422
    |||||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
    |||||
Db 423 GAAGATCGACCCCGAGTCATAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 482
    |||||
Qy 141 GlyGluLys 143
    |||||
Db 483 GGTGAGCAG 491
```

Search completed: December 15, 2002, 00:04:25  
Job time : 57.7555 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: December 11, 2002, 11:39:49 ; Search time 6.34401 Seconds  
(without alignments)  
506.931 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFDSYNFSLIACVAND.....RPKKIOTRRPEYTPHLS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues  
Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/PCT08\_PUBCOMB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	1043	100.0	198 10	US-09-782-953-6	Sequence 6, Appli
2	1043	100.0	198 10	US-09-782-953-7	Sequence 7, Appli
3	970.5	93.0	197 10	US-09-782-953-12	Sequence 12, Appl
4	970.5	93.0	197 10	US-09-782-953-13	Sequence 13, Appl
5	920.5	88.3	198 10	US-09-782-953-3	Sequence 3, Appli
6	920.5	88.3	198 10	US-09-782-953-4	Sequence 4, Appli
7	875	83.9	197 10	US-09-782-953-15	Sequence 15, Appl
8	875	83.9	197 10	US-09-782-953-16	Sequence 16, Appl
9	632.5	60.6	197 10	US-09-782-953-9	Sequence 9, Appli
10	632.5	60.6	197 10	US-09-782-953-10	Sequence 10, Appl
11	615.5	59.0	192 10	US-09-782-953-18	Sequence 18, Appl
12	615.5	59.0	192 10	US-09-782-953-19	Sequence 19, Appl
13	592.5	56.8	241 10	US-09-782-953-21	Sequence 21, Appl
14	592.5	56.8	241 10	US-09-782-953-22	Sequence 22, Appl
15	591.5	56.7	212 10	US-09-782-953-24	Sequence 24, Appl
16	591.5	56.7	212 10	US-09-782-953-25	Sequence 25, Appl
17	369.5	35.4	142 10	US-09-925-302-790	Sequence 790, App
18	284	27.2	58 10	US-09-864-761-43076	Sequence 35379, A
19	284	27.2	58 10	US-09-864-761-43076	Sequence 43076, A

20	271.5	26.0	56 10	US-09-864-761-34111	Sequence 34111, A
21	90.5	8.7	4019 10	US-09-738-973-425	Sequence 425, App
22	86.5	8.3	822 10	US-09-824-734-3	Sequence 3, Appli
23	81	7.8	2237 12	US-10-033-026-8	Sequence 8, Appli
24	81	7.8	2339 12	US-10-033-026-6	Sequence 6, Appli
25	81	7.8	2343 12	US-10-033-026-4	Sequence 4, Appli
26	79.5	7.6	559 10	US-09-854-549-7	Sequence 7, Appli
27	79.5	7.6	559 10	US-09-836-561-6	Sequence 6, Appli
28	77.5	7.4	864 10	US-09-883-096-2	Sequence 2, Appli
29	75.5	7.2	472 10	US-09-940-836A-2	Sequence 2, Appli
30	75	7.2	195 9	US-09-902-941-1888	Sequence 1888, Ap
31	75	7.2	362 9	US-09-349-755-32	Sequence 32, Appl
32	75	7.2	362 9	US-09-166-334-32	Sequence 32, Appl
33	75	7.2	362 10	US-09-350-206-32	Sequence 32, Appl
34	75	7.2	590 10	US-09-815-242-10812	Sequence 10812, A
35	75	7.2	1885 10	US-09-920-346-2	Sequence 2, Appli
36	74	7.1	311 10	US-09-819-252-2	Sequence 2, Appli
37	74	7.1	311 10	US-09-942-217-1061	Sequence 1061, Ap
38	74	7.1	311 10	US-09-833-263-1061	Sequence 1061, Ap
39	74	7.1	445 9	US-09-349-755-5	Sequence 5, Appli
40	74	7.1	445 9	US-09-166-334-5	Sequence 5, Appli
41	74	7.1	445 10	US-09-350-206-5	Sequence 5, Appli
42	74	7.1	916 10	US-09-745-763-174	Sequence 174, App
43	73	7.0	721 12	US-10-025-187-2	Sequence 2, Appli
44	72.5	7.0	1924 9	US-09-866-557A-2	Sequence 2, Appli
45	71.5	6.9	284 10	US-09-764-864-843	Sequence 843, App

ALIGNMENTS

RESULT 1  
US-09-782-953-6  
; Sequence 6, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: CALCI NEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-782-953-6

Query Match		100.0%	Score 1043;	DB 10;	Length 198;
Best Local Similarity		100.0%	Pred. No. 4,5e-98;		
Matches 198;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDRFD SYNFSLLIACVANDVFESETRAKESLFR	YDKD	TTQYFKSFKVRINFSN	60
DB	1	MDRFD SYNFSLLIACVANDVFESETRAKESLFR	YDKD	TTQYFKSFKVRINFSN	60
QY	61	PLSADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPNPDKQLISPPASPPVGVKQV	120		
DB	61	PLSADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPNPDKQLISPPASPPVGVKQV	120		
QY	121	EDATPVNVDLLYATSKLGPGEKYLHAATDTPSVVHVCS	DQENEE	EEEEEMERMKRP	180
DB	121	EDATPVNVDLLYATSKLGPGEKYLHAATDTPSVVHVCS	DQENEE	EEEEEMERMKRP	180
QY	181	KPKIOTRRPEYTPHLS	198		
DB	181	KPKIOTRRPEYTPHLS	198		

```
RESULT 2
US-09-782-953-7
; Sequence 7, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match          100.0%; Score 1043; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4,5e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
OY 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
OY 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
DB 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
OY 181 KPRTIOTRRPEYTPIHLS 198
DB 181 KPRTIOTRRPEYTPIHLS 198

RESULT 3
US-09-782-953-12
; Sequence 12, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 12
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-12

Query Match          93.0%; Score 970.5; DB 10; Length 197;
Best Local Similarity 93.4%; Pred. No. 9,2e-91;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

OY 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
```

```
OY 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
OY 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
DB 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
OY 181 KPRTIOTRRPEYTPIHLS 198
DB 181 KPRTIOTRRPEYTPIHLS 197

RESULT 4
US-09-782-953-13
; Sequence 13, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match          93.0%; Score 970.5; DB 10; Length 197;
Best Local Similarity 93.4%; Pred. No. 9,2e-91;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

OY 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MDFRFSYVNFSSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKRVIRINFSN 60
OY 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
DB 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPVGMKQV 120
OY 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
DB 121 EDATPVINVDLLVAISKLGPEKYEELHAATDPTPSVVVHVCSQDENEEBEEEMERMRKP 180
OY 181 KPRTIOTRRPEYTPIHLS 198
DB 181 KPRTIOTRRPEYTPIHLS 197

RESULT 5
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
```



Qy 121 EDATVINYDLVAISKLGPEKYEHLAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 180  
 Db 121 EDATVINYDLVAISKLGPEKYEHLAATDPTPSVVVHVCSDDQ-KEEBEEMERMRP 179  
 Qy 181 KPKIOTRRPEYTPIHLS 198  
 Db 180 KPKIOTRRPEYTPIHLS 197

# RESULT 9 US-09-782-953-9 ; Sequence 9, Application US/09782953 ; Patent No. US20020150953A1 ; GENERAL INFORMATION: ; APPLICANT: WILLIAMS, R. SANDERS ; APPLICANT: ROTHERMEL, BEVERLY ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP) ; FILE REFERENCE: US/09/782,953 ; CURRENT APPLICATION NUMBER: US/09/782,953 ; PRIOR FILING DATE: 2001-02-13 ; PRIOR APPLICATION NUMBER: 60/216,601 ; NUMBER OF SEQ ID NOS: 27 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 9 ; LENGTH: 197 ; TYPE: PRT ; ORGANISM: Mus musculus US-09-782-953-9

Query Match 60.6%; Score 632.5; DB 10; Length 197;  
 Best Local Similarity 64.4%; Pred. No. 1.2e-56;  
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

Qy 11 SGLIACVANDVFSSESETRAKFESLFRITYDKDTTFQYKSFKRVIRINFSNPLSADARLR 70  
 Db 11 STLVACVADVVEFTNOEVEKEKFGGLFRITDECVTQFLPKSFRRVRINFSHPSAARARIE 70  
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS-----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126  
 Db 71 LHEQFGRGKLLYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKQPISDATPV 130  
 Qy 127 INYDLVAISKLGPEKYEHLAATDPTPSVVVHVCSDDQNEEBEEMERMRKRPKPKIIQ 186  
 Db 121 LNYDLVAIAKLGPEKYEHLAAGTSTPSVVVHVCSDDMEBEDPK-----TSPKPKIIQ 185  
 Qy 187 TRRPEYTP 194  
 Db 186 TRRPEYTP 193

RESULT 10  
 US-09-782-953-10  
 ; Sequence 10, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: US/09/782,953  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus

## US-09-782-953-10

Query Match 60.6%; Score 632.5; DB 10; Length 197;  
 Best Local Similarity 64.4%; Pred. No. 1.2e-56;  
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

Qy 11 SGLIACVANDVFSSESETRAKFESLFRITYDKDTTFQYKSFKRVIRINFSNPLSADARLR 70  
 Db 11 STLVACVADVVEFTNOEVEKEKFGGLFRITDECVTQFLPKSFRRVRINFSHPSAARARIE 70  
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS-----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126  
 Db 71 LHEQFGRGKLLYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKQPISDATPV 130  
 Qy 127 INYDLVAISKLGPEKYEHLAATDPTPSVVVHVCSDDQNEEBEEMERMRKRPKPKIIQ 186  
 Db 121 LNYDLVAIAKLGPEKYEHLAAGTSTPSVVVHVCSDDMEBEDPK-----TSPKPKIIQ 185  
 Qy 187 TRRPEYTP 194  
 Db 186 TRRPEYTP 193

## RESULT 11 US-09-782-953-18 ; Sequence 18, Application US/09782953 ; Patent No. US20020150953A1 ; GENERAL INFORMATION: ; APPLICANT: WILLIAMS, R. SANDERS ; APPLICANT: ROTHERMEL, BEVERLY ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP) ; FILE REFERENCE: US/09/782,953 ; CURRENT APPLICATION NUMBER: US/09/782,953 ; PRIOR FILING DATE: 2001-02-13 ; PRIOR APPLICATION NUMBER: 60/216,601 ; NUMBER OF SEQ ID NOS: 27 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 18 ; LENGTH: 192 ; TYPE: PRT ; ORGANISM: Homo sapiens US-09-782-953-18

Query Match 59.0%; Score 615.5; DB 10; Length 192;  
 Best Local Similarity 63.3%; Pred. No. 6e-55;  
 Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

Qy 11 SGLIACVANDVFSSESETRAKFESLFRITYDKDTTFQYKSFKRVIRINFSNPLSADARLR 70  
 Db 6 STLVACVADVVEFTNOEVEKEKFGGLFRITDECVTQFLPKSFRRVRINFSHPSAARARIE 65  
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS-----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126  
 Db 66 LHEQFGRGKLLYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKQPISDATPV 125  
 Qy 127 INYDLVAISKLGPEKYEHLAATDPTPSVVVHVCSDDQNEEBEEMERMRKRPKPKIIQ 186  
 Db 126 LNYDLVAIAKLGPEKYEHLAAGTSTPSVVVHVCSDDMEBEDPK-----TSPKPKIIQ 180  
 Qy 187 TRRPEYTP 194  
 Db 181 TRRPEYTP 188

RESULT 12  
 US-09-782-953-19  
 ; Sequence 19, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: US/09-782,953  
; CURRENT APPLICATION NUMBER: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-19

Query Match 59.0%; Score 615.5; DB 10; Length 192;  
Best Local Similarity 63.3%; Pred. No. 6e-55;  
Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;  
QY 11 SLLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVIRNFNSPLSADARUR 70  
DB 6 STLVACVQDVVEFTNOEVEKGGFLRTYDDCVTFQYFKSFKRVIRNFNSPKSARARIE 65  
QY 71 LKTEFLGKEMKLYRAQTLHIGSS-----HLAPENPKQFLISPPASPPVGMKQVEDATPV 126  
DB 66 LHETPRGKGLKLYFAQVOTPETDQKHLAPPOPAKQFLISPPSPVSWQPINATPV 125  
QY 127 INYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMERMKRPKPKIIQ 186  
DB 126 LNYDLLYAVAKLPGKGYELHAGTSTPSVVHVHVCSDIEEDPK-----TSPKPKIIQ 180  
QY 187 TRPEYTP 194  
DB 181 TRPGLPP 188

RESULT 13  
US-09-782-953-21  
; Sequence 21, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: US/09-782,953  
; CURRENT APPLICATION NUMBER: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-21

Query Match 56.8%; Score 592.5; DB 10; Length 241;  
Best Local Similarity 61.5%; Pred. No. 1.7e-52;  
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;  
QY 1 MDRDFSYNFFSLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVIRNFNS 60  
DB 40 MDLSLP---TSLFACSVHEAVFEAREQERFEALTIYDDQVTFQYFKSFKRVIRNFNSK 96  
QY 61 PLSSADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHAPPNPDKQLISPPASPPVG 116  
DB 97 PEAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVG 156  
QY 117 WKQVEDATPVINYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMER 176  
DB 157 WKQSDAMPVINYDLLCAVSKLPGKGYELHAGTSTPSVVHVHVCSETEEBEE----- 210

QY 177 MKRPKPKIIQTRRPE 191  
DB 211 TKNPKQKIAQTRRPD 225  
RESULT 14  
US-09-782-953-22  
; Sequence 22, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: US/09-782,953  
; CURRENT APPLICATION NUMBER: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-22

Query Match 56.8%; Score 592.5; DB 10; Length 241;  
Best Local Similarity 61.5%; Pred. No. 1.7e-52;  
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;  
QY 1 MDRDFSYNFFSLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVIRNFNS 60  
DB 40 MDLSLP---TSLFACSVHEAVFEAREQERFEALTIYDDQVTFQYFKSFKRVIRNFNSK 96  
QY 61 PLSSADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHAPPNPDKQLISPPASPPVG 116  
DB 97 PEAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVG 156  
QY 117 WKQVEDATPVINYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMER 176  
DB 157 WKQSDAMPVINYDLLCAVSKLPGKGYELHAGTSTPSVVHVHVCSETEEBEE----- 210  
QY 177 MKRPKPKIIQTRRPE 191  
DB 211 TKNPKQKIAQTRRPD 225

RESULT 15  
US-09-782-953-24  
; Sequence 24, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: US/09-782,953  
; CURRENT APPLICATION NUMBER: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-24  
Query Match 56.7%; Score 591.5; DB 10; Length 212;  
Best Local Similarity 62.3%; Pred. No. 1.8e-52;

Search completed: December 11, 2002, 11:50:41  
Job time : 7.34401 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 136.563 Seconds  
(without alignments)  
934.784 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRDFSYNFSLLIACVAND.....RKPKEIQTTRRPTIHL 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US097 COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US099 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US60 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	198	21	US-09-782-953-6
2	1043	100.0	198	21	US-09-782-953-7
3	992.5	95.2	197	19	US-09-575-580B-8
4	970.5	93.0	197	21	US-09-782-953-12
5	970.5	93.0	197	21	US-09-782-953-13
6	970.5	93.0	197	26	US-10-247-671-174

7	970.5	93.0	197	27	US-60-323-784-174	Sequence 174, Appl
8	920.5	88.3	198	19	US-09-575-580B-4	Sequence 4, Appl
9	920.5	88.3	198	21	US-09-782-953-3	Sequence 3, Appl
10	920.5	88.3	198	21	US-09-782-953-4	Sequence 4, Appl
11	875	83.9	197	21	US-09-782-953-15	Sequence 15, Appl
12	875	83.9	197	21	US-09-782-953-16	Sequence 16, Appl
13	875	83.9	197	21	US-09-782-953-17	Sequence 17, Appl
14	843.5	80.9	197	19	US-10-104-047-2892	Sequence 18, Appl
15	632.5	60.6	197	21	US-09-575-580B-7	Sequence 19, Appl
16	632.5	60.6	197	21	US-09-782-953-9	Sequence 20, Appl
17	628.5	60.3	197	21	US-09-782-953-10	Sequence 21, Appl
18	628.5	60.3	197	21	US-09-782-953-11	Sequence 22, Appl
19	623	59.7	234	24	PCT-US02-17382-132	Sequence 23, Appl
20	623	59.7	234	24	US-10-030-613-1	Sequence 24, Appl
21	623	59.7	234	24	US-60-142-678-1	Sequence 25, Appl
22	623	59.7	234	24	US-10-104-047-2216	Sequence 26, Appl
23	615.5	59.0	192	20	US-09-614-474-2	Sequence 27, Appl
24	615.5	59.0	192	20	US-09-614-474-10	Sequence 28, Appl
25	615.5	59.0	192	21	US-09-782-953-18	Sequence 29, Appl
26	613.5	58.8	242	19	US-09-782-953-19	Sequence 30, Appl
27	592.5	56.8	241	1	US-09-575-580B-5	Sequence 31, Appl
28	592.5	56.8	241	1	PCT-US02-17382-133	Sequence 32, Appl
29	592.5	56.8	241	20	US-09-614-474-11	Sequence 33, Appl
30	592.5	56.8	241	21	US-09-782-953-21	Sequence 34, Appl
31	591.5	56.7	212	21	US-09-782-953-22	Sequence 35, Appl
32	591.5	56.7	212	21	US-09-782-953-24	Sequence 36, Appl
33	581	55.7	236	19	US-09-782-953-25	Sequence 37, Appl
34	432	41.4	142	21	US-09-575-580B-24	Sequence 38, Appl
35	432	41.4	142	21	US-09-782-953-26	Sequence 39, Appl
36	372	35.7	292	20	US-10-235-926-6466	Sequence 40, Appl
37	372	35.7	292	20	US-09-614-150-41193	Sequence 41, Appl
38	372	35.7	292	20	US-09-619-049-1440	Sequence 42, Appl
39	372	35.7	292	27	US-60-167-324-718	Sequence 43, Appl
40	372	35.7	292	27	US-60-171-627-2109	Sequence 44, Appl
41	372	35.7	292	27	US-60-173-386-680	Sequence 45, Appl
42	372	35.7	292	27	US-60-175-871-764	Sequence 46, Appl
43	372	35.7	292	27	US-60-184-775-692	Sequence 47, Appl
44	372	35.7	292	27	US-60-191-637-40818	Sequence 48, Appl
45	369.5	35.4	142	1	US-60-191-700-744	Sequence 49, Appl
					PCT-US00-05918-790	Sequence 50, Appl

#### ALIGNMENTS

RESULT 1  
US-09-782-953-6  
Sequence 6, Application US/09782953  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: UTS02:674821  
CURRENT APPLICATION NUMBER: US/09/782, 953  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/216,601  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-782-953-6

Query Match 100.0%; Score 1043; DB 21; Length 198;  
Best Local Similarity 100.0%; Pred. No. 4.1e-95;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDRDFSYNFSLLIACVANDVFSESETRAKFESLFYDYKDTTFQYKSKPRVIRNFSN 60  
Db 1 MDRDFSYNFSLLIACVANDVFSESETRAKFESLFYDYKDTTFQYKSKPRVIRNFSN 60

Qy	61	PLSADALRLHKTEFLGKEMKLTVEAOTLHIGSSHLAPNPDQOFLSPASPVGKQV	120
Db	61	PLSADALRLHKTEFLGKEMKLTVEAOTLHIGSSHLAPNPDQOFLSPASPVGKQV	120
Qy	121	BDATVINYDILVAISKGPGEKELHAATDPTESVVHVHVESDOENBEEBENMRKP	180
Db	121	BDATVINYDILVAISKGPGEKELHAATDPTESVVHVHVESDOENBEEBENMRKP	180
Qy	181	KPKIOTRRPEYTPHLS	198
Db	181	KPKIOTRRPEYTPHLS	198

```

RESULT 2
US-09-782-953-7
; Sequence 7, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCIENEURIN INTERACTING PROTEIN (WCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782, 953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

```

```

1 RESULT 3
2 US-09-575-5808-8
3 Sequence 8, Application US/09575580B
4 GENERAL INFORMATION:
5 APPLICANT: Mckean, P.
6 APPLICANT: Kayako, K.
7 APPLICANT: Ryeon, S.
8 TITLE OF INVENTION: CALCIIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
9 TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
10 FILE REFERENCE: HMV-048.01
11 CURRENT APPLICATION NUMBER: US/09/575,580B
12 CURRENT FILING DATE: 2000-05-22
13 NUMBER OF SEQ ID NOS: 45
14 SOFTWARE: PatentIn Ver. 2.1
15 SEQ ID NO 8
16 LENGTH: 197
17 TYPE: PRT
18 ORGANISM: Cricetus griseus

```

[illegible]

```

US-09-782-953-12
RESULT 4
US-09-782-953-12
; Sequence 12, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCIUMINERIN INTERACTING PROTEIN (NCIP)
; FILE REFERENCE: US/D:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-12

```

RESULT 5  
 US-09-782-953-13  
 ; Sequence 13, Application US/09782953  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHBERG, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

```
; FILE REFERENCE: UTSD.674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match          93.0%; Score 970.5; DB 21; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120

QY 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 179

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 6
US-10-247-671-174
; Sequence 174, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 174
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5511889CD1
US-10-247-671-174

Query Match          93.0%; Score 970.5; DB 26; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120

QY 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 179

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 7
US-60-323-784-174
; Sequence 174, Application US/60323784
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 P
; CURRENT APPLICATION NUMBER: US/60/323,784
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 174
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5511889CD1
US-60-323-784-174

Query Match          93.0%; Score 970.5; DB 27; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFTYDKDITTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAOTLHIGSSHLAPPNDKQFLISPPASPPVGKQV 120

QY 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKGYELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 179

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 8
US-09-575-580B-4
; Sequence 4, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKeon, F.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-575-580B-4

Query Match          88.3%; Score 920.5; DB 19; Length 198;
Best Local Similarity 89.9%; Pred. No. 6.7e-83;
```



US-09-782-953-16

; Sequence 16, Application US/09782953

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

; FILE REFERENCE: UTSD:674P21

; CURRENT APPLICATION NUMBER: US/09/782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 16

; LENGTH: 197

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-782-953-16

Query Match

Best Local Similarity 83.9%; Score 875; DB 21; Length 197;

Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

QY 1 MDRPDSYFNSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFSN 60

DB 4 VDLQDLP---SATIACHLDPRVFDGLCRAKFESLRTYDKDTTFQYFKSFKVRINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKOFLISPPASPPVGWKQV 120

DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKOFLISPPASPPVGWKQV 120

QY 121 EDATPVINYDLLYAIKSGKGYELHAATDTPSVVVHVCSQDQNEEBEEMERMRKP 180

DB 121 EDATPVINYDLLYAIKSGKGYELHAATDTPSVVVHVCSQDQNEEBEEMERMRKP 179

QY 181 KPILQTRRPEYTPHLS 198

DB 180 KPILQTRRPEYTPHLS 197

US-10-104-047-2892

; Sequence 2892, Application US/10104047

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: Novel full length cdna

; FILE REFERENCE: H1-A0105

; CURRENT APPLICATION NUMBER: US/10/104,047

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER:

; PRIOR FILING DATE:

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2892

; LENGTH: 252

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-104-047-2892

Query Match

Best Local Similarity 83.9%; Score 875; DB 25; Length 252;

Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

QY 1 MDRPDSYFNSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFSN 60

DB 59 VDLQDLP---SATIACHLDPRVFDGLCRAKFESLRTYDKDTTFQYFKSFKVRINFSN 115

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKOFLISPPASPPVGWKQV 120

DB 116 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKOFLISPPASPPVGWKQV 175

QY 121 EDATPVINYDLLYAIKSGKGYELHAATDTPSVVVHVCSQDQNEEBEEMERMRKP 180

US-09-782-953-7

; Sequence 7, Application US/09575580B

; GENERAL INFORMATION:

; APPLICANT: McKeon, F.

; APPLICANT: KAYAKO, K.

; APPLICANT: RYU, S.

; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN

; TITLE OF INVENTION: USES AND REAGENTS RELATED THEREO

; FILE REFERENCE: HMV-048.01

; CURRENT APPLICATION NUMBER: US/09/575,580B

; CURRENT FILING DATE: 2000-05-22

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-782-953-7

DB 176 EDATPVINYDLLYAIKSGKGYELHAATDTPSVVVHVCSQDQNEEBEEMERMRP 234

QY 181 KPILQTRRPEYTPHLS 198

DB 235 KPILQTRRPEYTPHLS 252

RESULT 14

US-09-575-580B-7

; Sequence 7, Application US/09575580B

; GENERAL INFORMATION:

; APPLICANT: McKeon, F.

; APPLICANT: KAYAKO, K.

; APPLICANT: RYU, S.

; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN

; TITLE OF INVENTION: USES AND REAGENTS RELATED THEREO

; FILE REFERENCE: HMV-048.01

; CURRENT APPLICATION NUMBER: US/09/575,580B

; CURRENT FILING DATE: 2000-05-22

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-575-580B-7

Query Match

Best Local Similarity 80.9%; Score 843.5; DB 19; Length 170;

Matches 161; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 30 AKFESLFRDYDKDTTFQYFKSFKVRINFSNPLSADARLRLHKTFLGKEMKLYPAQTL 89

DB 4 AKFESLFRDYDKDTTFQYFKSFKVRINFSNPLSADARLRLHKTFLGKEMKLYPAQTL 63

QY 90 HIGSSHLAPPNDPKOFLISPPASPPVGWKQVEDATPVINYDLLYAIKSGKGYELHAA 149

DB 64 HIGSSHLAPPNDPKOFLISPPASPPVGWKQVEDATPVINYDLLYAIKSGKGYELHAA 123

QY 150 TDTPSVVVHVCSQDQNEEBEEMERMRKPPIQTRRPEYTPHLS 197

DB 124 TDTPSVVVHVCSQDQNEEBEEMERMRKPPIQTRRPEYTPHLS 170

RESULT 15

US-09-782-953-9

; Sequence 9, Application US/09782953

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

; FILE REFERENCE: UTSD:674P21

; CURRENT APPLICATION NUMBER: US/09/782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 197

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-782-953-9

Query Match

Best Local Similarity 60.6%; Score 632.5; DB 21; Length 197;

Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 SSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFSNPLSADARL 70

DB 11 STLVAQVDVVEFTNQEVKEXFGLRTYDCVTLQFLKSFRRVRINFSHPKSAARIE 70



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 30.0506 Seconds  
(without alignments)  
2020.659 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFDSYNFSSLIACVAND.....RPKPKIIQTRRPEYPIHLS 198

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0782953/runat 11122002 114430 17377/app\_query.fasta\_1.1173  
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0782953 @CEN\_1\_1\_29 @runat 11122002 114430 17377 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	844.5	81.0	2174	2	US-08-665-040-1
2	94	9.0	1820	5	PCN-US94-09752-1
3	81	7.8	3138	3	US-09-234-332-5
4	81	7.8	3728	1	US-08-111-939-1
5	81	7.8	5467	2	US-07-745-206A-12
6	81	7.8	5467	2	US-08-311-363-12
7	81	7.8	7175	1	US-08-455-543A-8
8	81	7.8	7175	2	US-08-193-078B-8
9	81	7.8	7175	2	US-08-223-305C-8
10	81	7.8	7175	2	US-08-149-097D-8
11	81	7.8	7175	2	US-08-949-386-8
12	81	7.8	7175	3	US-08-450-562-8
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 5, Appli
					Sequence 12, Appli
					Sequence 12, Appli
					Sequence 8, Appli
					Sequence 8, Appli
					Sequence 8, Appli
					Sequence 8, Appli
					Sequence 8, Appli

13	81	7.8	7175	4	US-08-984-709A-8	Sequence 8, Appli
14	81	7.8	7175	4	US-08-450-272-8	Sequence 8, Appli
15	81	7.8	7177	4	US-09-268-163-7	Sequence 7, Appli
16	81	7.8	7362	1	US-08-455-543A-7	Sequence 7, Appli
17	81	7.8	7362	2	US-08-193-078B-7	Sequence 7, Appli
18	81	7.8	7362	2	US-08-223-305C-7	Sequence 7, Appli
19	81	7.8	7362	2	US-08-149-097D-7	Sequence 7, Appli
20	81	7.8	7362	3	US-08-949-386-7	Sequence 7, Appli
21	81	7.8	7362	3	US-08-450-562-7	Sequence 7, Appli
22	81	7.8	7362	4	US-08-984-709A-7	Sequence 7, Appli
23	81	7.8	7362	4	US-08-450-272-7	Sequence 7, Appli
24	81	7.8	7364	4	US-09-268-163-5	Sequence 5, Appli
25	81	7.8	7376	4	US-09-268-163-3	Sequence 3, Appli
26	80.5	7.7	6232	4	US-08-456-200B-11	Sequence 11, Appli
27	80	7.7	1113	1	US-08-231-342-5	Sequence 5, Appli
28	80	7.7	3093	1	US-08-252-966B-19	Sequence 19, Appli
29	79	7.6	1113	1	US-08-231-342-7	Sequence 7, Appli
30	79	7.6	36651	4	US-09-738-894A-3	Sequence 3, Appli
31	78.5	7.5	1644	4	US-09-877-730-25	Sequence 25, Appli
32	78.5	7.5	1875	4	US-09-877-730-23	Sequence 23, Appli
33	78.5	7.5	1887	4	US-09-877-730-29	Sequence 29, Appli
34	78.5	7.5	2139	4	US-09-877-730-21	Sequence 21, Appli
35	78.5	7.5	2382	4	US-09-877-730-27	Sequence 27, Appli
36	78.5	7.5	2481	4	US-09-877-730-15	Sequence 15, Appli
37	78.5	7.5	2715	4	US-09-877-730-5	Sequence 5, Appli
38	78.5	7.5	2724	4	US-09-877-730-19	Sequence 19, Appli
39	78.5	7.5	2958	4	US-09-877-730-9	Sequence 9, Appli
40	78.5	7.5	2976	4	US-09-877-730-11	Sequence 11, Appli
41	78.5	7.5	3210	4	US-09-877-730-17	Sequence 17, Appli
42	78.5	7.5	3219	4	US-09-877-730-7	Sequence 7, Appli
43	78.5	7.5	3453	4	US-09-877-730-31	Sequence 31, Appli
44	78.5	7.5	3874	4	US-09-877-730-7	Sequence 1, Appli
45	78.5	7.5	6171	1	US-08-459-568-1	

## ALIGNMENTS

RESULT 1  
US-08-665-040-1  
; Sequence 1, Application US/08665040  
; Patent No. 5869118  
; GENERAL INFORMATION:  
; APPLICANT: ESTIVILL PALLEJA, XAVIER  
; APPLICANT: FUENTES, JUAN JOSE  
; APPLICANT: PRITCHARD, MELANIE  
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,  
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
; TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY  
; STREET: 26 WEST 61ST STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10023  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT 5.1 FOR DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,040  
; FILING DATE: JUNE 7, 1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: ES P9501140  
; FILING DATE: JUNE 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JANET I. CORD

REGISTRATION NUMBER: 33,778  
REFERENCE/DOCKET NUMBER: U010815-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800  
TELEFAX: (212) 246-8959  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: cDNA for mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: human  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: foetal  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE: gene library of cDNA  
LIBRARY: gene library of cDNA from foetal  
CLONE: BC-17.8-1 and BC-17.8-2  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9  
MAP POSITION: 21q22.1-q22.2  
FEATURE:  
NAME/KEY: cDNA for mRNA, BC-17.8  
LOCATION: 1..2174  
FEATURE:  
NAME/KEY: untranslated 5'  
LOCATION: 1..48  
FEATURE:  
NAME/KEY: coding sequence  
LOCATION: 49..560  
OTHER INFORMATION: Down Syndrome critical  
region 1 (DSCR1)  
FEATURE:  
NAME/KEY: DSCR1  
LOCATION: 1..171 PEPTIDES  
IDENTIFICATION METHOD: translation of the  
OTHER INFORMATION: DSCR1 sequence.  
OTHER INFORMATION: - deduced protein  
OTHER INFORMATION: - proline-rich protein domains  
OTHER INFORMATION: - glutamic acid-rich protein domains  
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains  
FEATURE:  
NAME/KEY: untranslated 3'  
LOCATION: 564..2174  
FEATURE:  
NAME/KEY: 2 poly (A)  
LOCATION: 1541..1546 AND 2132..2137  
US-08-665-040-1  
Alignment Scores:  
Pred. No.: 1.95e-98 Length: 2174  
Score: 844.50 Matches: 161  
Percent Similarity: 97.63% Conservative: 4  
Best Local Similarity: 95.27% Mismatches: 3  
Query Match: 80.97% Indels: 1  
Gaps: 2  
US-09-782-953-7 (1-198) x US-08-665-040-1 (1-2174)  
QY 30 AlAlaYsPhcglUSeRLeuPheArGThrTYrAspLYsApThrThrPhcglInTyRPhelys 49  
DB 58 GCCAAATTGAGTCCCTTCTTAGACGATGACAGACATCACTTTCAGTATTTAAAG 117  
QY 50 SerPheLYsArgValAlaGllLeaenPhSeRAsnProLeuSerAlaAlaAspAlaArgLeu 69  
DB 118 AGCTTCAACAGAGTCAGAAATTAACCTCAGCAACCCCTTCCCGAGAGATGCCAGGCTC 177  
QY 70 ArgLeuHIsLYeRThRglUpheLeuGllYleaglUmetLYsLeuTYrPheAlaGlnThrLeu 89  
DB 178 CAGCTGATTAAGACTGAGTTCCTGCGAAAGGAATGAAGTTATTTTGTCTCAGACCTTA 237

QY 90 HisIleGlySerSerHisIleAlaAProProAsnProAspLYsGlnPheLeuIleSerPro 109  
DB 238 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGAGTTCTGATCTCCCT 297  
QY 110 ProIleSerProProValAlaGlyTrpLYsGlnValAlaAspAlaThrProValIleAsnTYR 129  
DB 298 CCCGCTCTCCGACGATGGGATGGAAACAACTGAAAGATCCGACCCGACGATAACTAT 357  
QY 130 AspleuLeuTYrAlaIleSerLYsLeuGlyProGlyGluLYsTYrGluLeuHisAlaAla 149  
DB 358 GATCTTTATATGATCCATCTCCAGCTGGCCGAGGGGAAAGTATGATGACGACGAGCG 417  
QY 150 ThrAspProThrProSerValValIleValCysgluSerAspGlnGluAangGlu 169  
DB 418 ACTGACACCACTCCACGCTGGTGTCTCATGTATGTAGAGTGAAGGAGGAGGAG 474  
QY 170 GluGluGluGluMetGluAArgMetLYsArgProLYsProLYsIleIleGlnThrArgArg 189  
DB 475 GAAAGAGGAAATGAAAGAAATAGAGAGACCTTAAGCCAAATTAATTCAGACGAGAGG 534  
QY 190 ProGluTYrThrProIleHisLeuSer 198  
DB 535 CCGAGGTACAGCCGATCCACCTCAGC 561  
RESULT 2  
PCT-US94-09752-1  
Sequence 1, Application PC/TUS9409752  
GENERAL INFORMATION:  
APPLICANT: David S. Strayer and Avinash Chander  
TITLE OF INVENTION: Compositions and Methods for  
Targeting Cells and Modulating Pulmonary Surfactant Secretion  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jane Massey Licata, Esq.  
STREET: 210 Lake Drive East, Suite 201  
CITY: Cherry Hill  
STATE: NJ  
COUNTRY: USA  
ZIP: 08002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB  
MEDIUM TYPE: STORAGE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09752  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/176,218  
FILING DATE: December 30, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/114,951  
FILING DATE: August 31, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: JEFF-0042  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 779-2400  
TELEFAX: (609) 779-8488  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1820  
TYPE: nucleic acid  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
PCT-US94-09752-1  
Alignment Scores:  
Pred. No.: 0.0363 Length: 1820  
Score: 94.00 Matches: 49



**Alignment Scores:**

## REFERENCE/DOCK

REGISTRATION NUMBER: 33,694  
REFERENCE/DOCKET NUMBER: 02481.1321-00000  
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
STRAIN: osteoblastic cell line MC3T3E1
FEATURE:
NAME/KEY: CDS
LOCATION: 69..3452
US-08-111-939-1

Alignment Scores:
Pred. No.: 5.13 Length: 3728
Score: 81.00 Matches: 33
Percent Similarity: 38.99% Conservative: 29
Best Local Similarity: 20.75% Mismatches: 27
Query Match: 7.77% Indels: 60
DB: 1 Gaps: 8

US-09-782-953-7 (1-198) x US-08-111-939-1 (1-3728)

QY 50 SerPheLysArgValArg-----11e 56
DB 2997 AACTGGAGACGCGATTCGGGAGATCTTGCTATGACGGAGACCGTCCCTTCGAGTT 3056
QY 57 AenPheSerAnPheLysSerAlaAlaPheAlaArgLeuArgLeuHisIleYsThrGluPhe 76
DB 3057 GACCCCTCAGACACCTGAGACCCCGGACGCGCGCATGAGAGCGCGCTACAGTAC 3116
QY 77 -----LeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHisIleGlySer 93
DB 3117 CGGCTCCGATGAGGAGGAGATGCAACTG-----CGTGGCTCAATCTTACCGCA 3167
QY 94 SerHisLeuAlaProPheAnPheProAlaPheGlnPheLeuLysSerProPheAlaSerPro 113
DB 3168 GGCCCTGCCCAAGGCCCACTCT-----GCCCTTATGCTCTCCCTTCCCT 3215
QY 114 -----ProValGlyTyrPhe 118
DB 3216 ACACGACCATTAACCTTGAAGCCCTGGGAATTCTACCACTACCACTGAGCGCTGGAG 3275
QY 119 -----GlnValGluPheAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSer 136
DB 3276 GAGTCAGAGACTGACCTTACGAGAGTAGTGACA-----3311
QY 137 LysLeuGlyProGlyGluLysTyrGluLeuHisIleAlaThrAspProThrProSerVal 156
DB 3312 -----GAGTTGACACAGATAGGACTGAC-----3338
QY 157 ValValHisValCysGluSerAspGlnGluAsnGluGluGluGluGluMetGlu 175
DB 3339 ---CTAAGAGTGGAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGCAC 3392

RESULT 5
US-07-745-206A-12
Sequence 12, Application US/07745206A
Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392, 3396..3489, 3495..3539, 3543..3581, 3585
LOCATION: ..3587, 3591..3626, 3630..3689, 3737..3744
LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12

Alignment Scores:
Pred. No.: 9.26 Length: 5467
Score: 81.00 Matches: 50
Percent Similarity: 33.77% Conservative: 27
Best Local Similarity: 21.93% Mismatches: 80
Query Match: 7.77% Indels: 71
DB: 1 Gaps: 9

US-09-782-953-7 (1-198) x US-07-745-206A-12 (1-5467)

QY 7 SerTyrAnPheSerSerLeuIleAlaCysValAlaAnPhePheValPheSerGluSer 26
DB 2232 AACTACACTGCTGTAAGATGCTTCTGCGCATGCTGTGCAACATGGCCAGCCCA 2291
QY 27 GlnThrAlaGlnAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46
DB 2292 GAGCTGACCAAGATGAAAGAGATGGAAGAGACAGCCATCCAGACTTGTGCA 2351
QY 47 TyrPheLysSerPheLysArgValArgIleAnPheSerAnProLeuSerAlaAlaAsp 66
DB 2352 AAGCCCAAGAGATGCTGAGATC-----AGCCCATGTGCGCGCAAC 2396
QY 66 -----66
DB 2397 ATCTCATGCGCCGACGACGAGAACTCGGCAAGGCGCGTGTGGAGGAGCGG 2456
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
DB 2457 GCCAGCAGCTACGGCTGAGAACTGGGGGCCAGCTGCGAGCCCTGTACAGCGAGATG 2516
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
DB 2517 GACCCGAGAGAGCGGTGGCTTGGCACTAGCGCCACCTG-----2558
QY 98 ProPheAnPheAspLysGlnPheLeuLysSerProPro-----110
DB 2559 -----CGGCCGACATGAGAGAGCACTGACCGGCGCTGTGTGTGAGAGCTGGCGCC 2612

```



TITLE OF INVENTION: METHODS  
 NUMBER OF SEQUENCES: 57  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Brown, Martin, Haller & McClain  
 STREET: 1660 Union Street  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92101-2926  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/455,543A  
 FILING DATE: May 31, 1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/223,305  
 FILING DATE: April 4, 1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/868,354  
 FILING DATE: April 10, 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/745,206  
 FILING DATE: 15-AUG-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/620,250  
 FILING DATE: 30-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/482,384  
 FILING DATE: 20-FEB-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/603,751  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US89/01408  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/176,899  
 FILING DATE: 04-APR-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 6362-52517  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619)238-0999  
 TELEFAX: (619)238-0062  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7175 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 144..6857  
 NAME/KEY: 5' UTR  
 LOCATION: 1..143  
 FEATURE:  
 NAME/KEY: 3' UTR  
 LOCATION: 6855..7175  
 US-08-455-543A-8

Alignment Scores:  
 Pred. No.: 14.1  
 Score: 81.00  
 Percent Similarity: 33.79%  
 Best Local Similarity: 21.46%  
 Query Match: 7.77%  
 DB: 1

Length: 7175  
 Matches: 47  
 Conservative: 27  
 Mismatches: 75  
 Indels: 70  
 Gaps: 9

US-09-782-953-7 (1-198) x US-08-455-543A-8 (1-7175)  
 QY 7 SerTyrAsnPhseSerLeuIleAlaCysValAlaAsnAspAspValPheSerGluSer 26  
 Db 2232 AACCTACACTCTGCTGAATGCTTTCTTGCGCCATCGCTGGACAACTGGACCAACCCCA 2291  
 QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46  
 Db 2292 GACCTGACCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2351  
 QY 47 TyrPheLysSerPheLysArgValAlaGlnLeuPheSerAspProLeuSerAlaAlaAsp 66  
 Db 2352 AAGGCCAAGAGTGGCTGAATC-----AGCCCAATGCTGCTCCCGGAAC 2396  
 QY 66 ----- 66  
 Db 2397 ATCTCATGCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2456  
 QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGluLysGluMet 81  
 Db 2457 GCCAGCCAGCTACGCGCTGACAACTGCGGCGCACTGCGAGCGCTGTACACCGAGATG 2516  
 QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97  
 Db 2517 GACCCCGAGAGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 2558  
 QY 98 ProProAsnProAspLysGlnPheLeuLysSerProPro----- 110  
 Db 2559 -----CGGCCCGACATGAGACGACGACGACGACGACGACGACGACGACGACGACGAC 2612  
 QY 111 -----AlaSerProProValGlyTyrPheGlnValGlnAspAlaThrProValIleAsn 128  
 Db 2613 GACG 2663  
 QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyLysLysArgTyrGluLeuHisAla 148  
 Db 2664 -----GCCCGGAGGGGTGACCTCGCGGACGACGACGACGACGACGACGACGACG 2708  
 QY 149 AlaThrAspProThrProSerValValAlaHisValLysGluSerAspGlnGlu----- 166  
 Db 2709 GACAAGACCAAGACCCCGCGCGCG-----GGGACACAGAGACCGACGCA 2750  
 QY 167 -----AangLugluGluGluGluGluMetGluArgMetLysArgProLysPro 182  
 Db 2751 GAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2807

RESULT 8  
 US-08-193-078B-8  
 Sequence 8, Application US/08193078B  
 Patent No. 5846757  
 GENERAL INFORMATION:  
 APPLICANT: Harpold, Michael  
 APPLICANT: Ellis, Steven  
 APPLICANT: Williams, Mark  
 APPLICANT: Feldman, Daniel  
 APPLICANT: McCue, Ann  
 APPLICANT: Brenner, Robert  
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
 NUMBER OF SEQUENCES: 29  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN  
 STREET: 1660 UNION STREET  
 CITY: SAN DIEGO  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92101  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,078B
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
; US-08-193-078B-8

```

## Alignment Scores:

```

Pred. NO.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 2 Gaps: 9

```

US-09-782-953-7 (1-198) x US-08-193-078B-8 (1-7175)

```

Qy 7 serTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGATGATGCTTTCTGCGCATCTGTGGACAACTGGCCAAACGCCAA 2291
Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGAGATGGAAGACGACCAATCAGAAGCTTGTCTGCAA 2351
Qy 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGGCCAAGAGTGGCTGAATC-----AGCCCCATGCTGCGCGAAC 2396
Qy 66 ----- 66
Db 2397 ATCTCCATCCGCCAGGACGAGAACTCGGCCAAGCGCGCTGCTGGAGCAGCGG 2456
Qy 67 -----AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTAGCGGTGACAGACCTGCGGGCCAGCTGCGGCGCTGTACAGCGAGATG 2516
Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGAGCGGTGGCGCTTCGCCACTACGCGCCACTG----- 2558
Qy 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
Db 2559 -----CGGCCCCGACATGAAGACGACCTGGACCGCCGCTGGTGGTGGAGCTGGCGCGC 2612

```

```

Qy 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGCGGGCGCGCTGGGAGGCAAGCCGACCTGAGGCTCGGAG----- 2663
Qy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGGGCGTGCACCTCCCGCGCAGGACCCACCGGACCGC 2708
Qy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2709 GACAAGGACAAAGACCCCGCGCG-----GGGACCCAGGACCGACGCA 2750
Qy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCGCCGAGGCGGAGCGGGAGCCCGTGCCTGGGAGGCGCGCGCGCG 2807

RESULT 9
US-08-223-305C-8
; Sequence 8, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062

```

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-223-305C-8  
Alignment Scores:  
Pred. No.: 14.1 Length: 7175  
Score: 81.00 Matches: 47  
Percent Similarity: 33.79% Conservative: 27  
Best Local Similarity: 21.46% Mismatches: 75  
Query Match: 7.77% Indels: 70  
Gaps: 9  
DB:  
US-09-782-953-7 (1-198) x US-08-223-305C-8 (1-7175)  
QY 7 SerTAAAnPheSerSerLeuIleAlaCyValAlaAsnAPhAPhValPheSerGluSer 26  
DB 2232 AACTACACTGTGTAATGCTTCTGGCCATGCTGAGACAACTGCGCAAGCCCAA 2291  
QY 27 GluThrArgAlaIlePheGluSerLeuPheArgThrTyrAspIleAspThrPheGln 46  
DB 2292 GAGCTGACCAAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2351  
QY 47 TyrPheIleSerPheIleValArgIleAsnPheSerAsnProIleSerAlaIleAsp 66  
DB 2352 AAGGCCAAAGAAATGCTGTAAGTC-----AGCCCAATGCTGCGCGCAAC 2396  
QY 66 ----- 66  
DB 2397 ATCTCATCCGCCCGACAGACAGAACTCGCCAGAGCGCGCTGCTGGAGCGCG 2456  
QY 67 ---AlaArgIleuArgLeuHis-----LysThrGluPheLeuGlyIleGluMet 81  
DB 2457 GCCAGCGAGCTAGCGGCTGCGAGAACTCGCGCGCGCGCGCGCTGACAGAGATG 2516  
QY 82 -----LysLeuTyrPheAlaGlnThrIleuHisIleGlySerSerHisLeuAla 97  
DB 2517 GACCCCGAGAGCGGCTGCGCTGCGCACTACGCGCACCTG----- 2558  
QY 98 ProProAsnProAspIleGlnPheLeuIleSerProPro----- 110  
DB 2559 -----CGCCCGCATGATGAAGACGACCTGAGCGCGCGCTGCTGCTGAGAGCGCGC 2612  
QY 111 -----AlaSerProProValGlyTyrIleGlnIleValAlaIleAsn 128  
DB 2613 GACGCGCGCGCGCGCGCTGCGAGCAAGCCGACCTGAGCGCTGCGAG----- 2663  
QY 129 TyrAspLeuLeuTyrAlaIleSerIleLeuGlyProGlyIleuIleValAla 148  
DB 2664 -----GCCCGCGAGCGCGCTGAGCCCTCGCGCAGACACACCGCGCACCGC 2708  
QY 149 AlaThrAspProThrProSerValValAlaHisValCysGluSerAspGlnIle----- 166  
DB 2709 GACAGAGACAAAGCCCGCGCG-----GGGACCGAGAGCGAGCA 2750  
QY 167 -----AsnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 182  
DB 2751 GAGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2807  
RESULT 10

US-08-149-097D-8  
Sequence 8, Application US/08149097D  
Patent No. 5874236  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
City: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,097D  
FILING DATE: 05-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/105,536  
FILING DATE: 11-AUG-1993  
PRIOR APPLICATION DATA: WO PCT/US92/06903  
APPLICATION NUMBER: 14-AUG-1992  
FILING DATE: 14-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/914,231  
FILING DATE: 13-JUL-1992  
PRIOR APPLICATION DATA: US 07/868,354  
APPLICATION NUMBER: 10-APR-1992  
FILING DATE: 10-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-55038  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

```

; FEATURE: CDS
; NAME/KEY: 144..6857
; LOCATION: 144..6857
; FEATURE: 5'UTR
; NAME/KEY: 1..143
; LOCATION: 1..143
; FEATURE: 3'UTR
; NAME/KEY: 6855..7175
; LOCATION: 6855..7175
US-08-149-097D-8

```

## Alignment Scores:

```

Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 2 Gaps: 9

```

US-09-782-953-7 (1-198) x US-08-149-097D-8 (1-7175)

```

QY 7 SerTyrAsnPheSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGGAATGTCTTTCTGCCATCGCTGGACAACCTGGCCACGCCAA 2291
QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGATGAAGAGCAGCCATCAGAACGCTTGCTCTCAA 2351
QY 47 TyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGCCAAAGAGTGGCTGAAGTC-----AGCCCATGTCTGCCCGCAG 2396
QY 66 ----- 66
Db 2397 ATCTCCATCGCCGACGACGACAGAACTCGGCCAAGCGCGCTCGGTGGAGACGCG 2456
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGTACGCTTCGAGAACTCGGGGCCAGCTCGCGGCGCTGTACAGCGAGATG 2516
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisLeuGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGGAGCGCTCGCTTCGCCACTACGCGCCACCTG----- 2558
QY 98 ProProAsnProAspLysGlnPheLeuLeuSerPro----- 110
Db 2559 -----CGGCCCGCATGAGAGCGCACCTCGACCGCGCGCTGGTGGAGTGGCGCGC 2612
QY 111 -----AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGGGGCGCGCTGGAGGAGCAAGCCGACCTGAGCTCGGAG----- 2663
QY 129 TyrAspLeuLeuTyrAlaLysSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGGGCGCTCGACCTCGCGCGAGCGCACCCACCGCACCGC 2708
QY 149 AlaThrAspProThrProSerValValHisValCysLeuSerAspGln----- 166
Db 2709 GACAGGACAGACCCCGCGCGG-----GGGACCGAGCAGCAGCA 2750
QY 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCGCCGAGGAGCGGAGCGCGGAGCGCGGTGCCCGGAGGAGCGCGCGCGCG 2807

```

## RESULT 11

US-08-949-386-8

; Sequence 8, Application US/08949386

; Patent No. 6090623

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

```

; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: US
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,386
; FILING DATE:

```

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/290,012

FILING DATE: 11-AUG-1994

APPLICATION NUMBER: 08/149,097

FILING DATE: 5-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/105,536

FILING DATE: 11-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 519808

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 238-0999

TELEFAX: (619) 238-0062

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 7175 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 144..6857

FEATURE:

NAME/KEY: 5'UTR

LOCATION: 1..143

FEATURE:

NAME/KEY: 3'UTR

LOCATION: 6855..7175

US-08-949-386-8

## Alignment Scores:

```

Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 3 Gaps: 9

```

US-09-782-953-7 (1-198) x US-08-949-386-8 (1-7175)

```

QY 7 SerTyrAsnPheSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGGAATGTCTTTCTGCCATCGCTGGACAACCTGGCCACGCCAA 2291
QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGATGAAGAGCAGCCATCAGAACGCTTGCTCTCAA 2351
QY 47 TyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsnProLeuSerAlaAlaAsp 66

```

Db 2352 AAGCCAAAGAGTGGCTGAAGTC-----ACCCCATCTCTGCGCGAAG 2396  
OY 66 ----- 66  
Db 2397 ATCTTCATCGCGCGAGGAGAGAACTCGGCAAGGCGCTGCTGGAGCAGCGG 2456  
OY 67 ---AlaATgLeuATgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81  
Db 2457 GCCAGCCAGCTACCGCTCCAGAACTCGGCGCCAGCTCGAGGCGCTGTACAGCGAGATG 2516  
OY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97  
Db 2517 GACCCCGAGAGAGCGGCTGCGCTTCCCACTACGCGCCACCTG----- 2558  
OY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110  
Db 2559 -----CGGCCCGCATGTAGAGACGCACTGACCGCGCGCTGTGTGTGAGCTGGCGCGC 2612  
OY 111 -----AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsn 128  
Db 2613 GACGCGCGCGCGCGCGCTGTGAGGCAAGCCCGACCTGAGGCTGCGGAG----- 2663  
OY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148  
Db 2664 -----GCCCCGAGGCGCTGCACTCGCGCGAGCAGACCGCGCGCGC 2708  
OY 149 AlaThrAspProThrProSerValValAlaHisValCysGluSerAspGlnLys----- 166  
Db 2709 GACACAGACACACACCGCGCGCGG-----GGGACACAGACAGACAGCA 2750  
OY 167 -----AsnGluGluGluGluGluGluMetLysMetLysProLysPro 182  
Db 2751 GAGGCGCGCGAGGCGAGAGCGGCGGAGCGCGCTGCGCGGAGAGCGCGCGCGCG 2807

RESULT 12  
US-08-450-562-8  
Sequence 8, Application US/08450562  
Patent No. 6096514  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: McCue, Ann  
APPLICANT: Gillespie, Alison  
APPLICANT: Feldman, Daniel  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: US  
ZIP: 92101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.35  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,562  
FILING DATE: 435  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/404,950  
FILING DATE: 13-MAR-1995  
APPLICATION NUMBER: 08/336,257  
FILING DATE: 7-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/314,083

FILING DATE: 28-SEPT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/311,363  
FILING DATE: 23-SEPT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290,012  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/223,305  
FILING DATE: 4-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/193,078  
FILING DATE: 07-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,097  
FILING DATE: 5-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/105,536  
FILING DATE: 11-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,231  
FILING DATE: 13-JULY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: 10-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06903  
FILING DATE: 14-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/603,751  
FILING DATE: 08-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/482,384  
FILING DATE: 02-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/176,899  
FILING DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-519812  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-450-562-8

Alignment Scores:  
Pred. No.: 14.1  
Length: 7175



Score: 81.00 Matches: 47  
 Percent Similarity: 33.79% Conservative: 27  
 Best Local Similarity: 21.46% Mismatches: 75  
 Query Match: 7.77% Indels: 70  
 DB: 3 Gaps: 9

US-09-782-953-7 (1-198) x US-08-450-562-8 (1-7175)

Qy 7 SerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26  
 Db 2232 AACTACACTCTGCTGAATGCTTTCTGCCATCGCTGGACACCTGGCCAAAGCCCA 2291  
 Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46  
 Db 2292 GAGTGTGACCAAGGATGAAGAGGAGATGAAGAAGCAGCAATCAGAGCTTGCTCTGCAA 2351  
 Qy 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66  
 Db 2352 AAGCCCAAGAGTGGCTGAAGTC-----AGCCCCATGTCTGCCGCGAAC 2396  
 Qy 66 ----- 66  
 Db 2397 ATCTCCATCGCCGAGCAGCAGAACTCGGCCAAGCGCGCTCGGTGGGACGCGG 2456  
 Qy 67 ---AlaArgIleuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81  
 Db 2457 GCCAGCCAGCTACGCTGCTGAGAACCTCGGGCCAGCTGCGAGCGCTGTACAGCGAGATG 2516  
 Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97  
 Db 2517 GACCCCGAGGAGCGCTGCGCTTCCCACTACGCCGACCTG----- 2558  
 Qy 98 ProAsnProAspLysGlnPheLeuLeuSerPro----- 110  
 Db 2559 -----CGGCCGACATGAAGACGACCTGGACCGCGCTGGTGGAGCTGGCGCG 2612  
 Qy 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128  
 Db 2613 GACG 2663  
 Qy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148  
 Db 2664 -----GCCCGCGAGGCGCTGACCTCGCGCGAGGACACCGCGCGCGCGCG 2708  
 Qy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166  
 Db 2709 GACAAGGACAAAGACCCCGCGCGCG-----GGGACACGAGCGCGGCA 2750  
 Qy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182  
 Db 2751 GAGCCCGGAGGAGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 2807

## RESULT 13

US-08-984-709A-8  
 ; Sequence 8, Application US/08984709A  
 ; Patent No. 6320032

## GENERAL INFORMATION:

APPLICANT: Williams, Mark E.  
 APPLICANT: Stauderman, Kenneth A.  
 APPLICANT: Harpold, Michael M.  
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
 TITLE OF INVENTION: METHODS  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:

ADDRESS: Heller Ehrman White & McAuliffe  
 STREET: 4250 Executive Square, Suite 700  
 CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/984,709A  
 FILING DATE: 02-DEC-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 450-8400  
 TELEFAX: (619) 587-5360  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7175 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:

NAME/KEY: CDS  
 LOCATION: 144..6857  
 FEATURE:

NAME/KEY: 5'UTR  
 LOCATION: 1..143  
 FEATURE:

NAME/KEY: 3'UTR  
 LOCATION: 6855..7175  
 US-08-984-709A-8

Alignment Scores:  
 Pred. No.: 14.1 Length: 7175  
 Score: 81.00 Matches: 47  
 Percent Similarity: 33.79% Conservative: 27  
 Best Local Similarity: 21.46% Mismatches: 75  
 Query Match: 7.77% Indels: 70  
 DB: 3 Gaps: 9

US-09-782-953-7 (1-198) x US-08-984-709A-8 (1-7175)

Qy 7 SerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26  
 Db 2232 AACTACACTCTGCTGAATGCTTTCTGCCATCGCTGGACACCTGGCCAAAGCCCA 2291  
 Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46  
 Db 2292 GAGTGTGACCAAGGATGAAGAGGAGATGAAGAAGCAGCAATCAGAGCTTGCTCTGCAA 2351  
 Qy 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66  
 Db 2352 AAGCCCAAGAGTGGCTGAAGTC-----AGCCCCATGTCTGCCGCGAAC 2396  
 Qy 66 ----- 66  
 Db 2397 ATCTCCATCGCCGAGCAGCAGAACTCGGCCAAGCGCGCTCGGTGGGACGCGG 2456  
 Qy 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81  
 Db 2457 GCCAGCCAGCTACGCTGCTGAGAACCTCGGGCCAGCTGCGAGCGCTGTACAGCGAGATG 2516  
 Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97  
 Db 2517 GACCCCGAGGAGCGCTGCGCTTCCCACTACGCCGACCTG----- 2558  
 Qy 98 ProAsnProAspLysGlnPheLeuLeuSerPro----- 110  
 Db 2559 -----CGGCCGACATGAAGACGACCTGGACCGCGCTGGTGGAGCTGGCGCG 2612  
 Qy 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128  
 Db 2613 GACG 2663

QY 129 TyraPleuLeuTyraAlaIleSerLysLeuGlyProGlyGlyTyraGluLeuHisAla 148  
Db 2664 -----GCCCGGAGGGCGTGCACCCCTCGCGCAGGACACCCGACCGC 2708  
QY 149 AlaThrAapProThiProSerValValHisValCysGluSerAapGlnGlu----- 166  
Db 2709 GACAAGACACAGACCCCGCGCGC-----GGGGACGAGGACCGGACGA 2750  
QY 167 -----AaGlnGluGlnGluGlnGluMetGluArgMetLysArgProLysPro 182  
Db 2751 GAGCGCCGAGGCGGAGAGCGGGAGCCCGGTGCGCCGAGAGACCGCGCGC 2807  
RESULT 14  
US-08-450-272-8  
Sequence 8, Application US/08450272  
Patent No. 6387696  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: McGue, Ann  
APPLICANT: Gillespie, Allison  
APPLICANT: Feldman, Daniel  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
METHODS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: US  
ZIP: 92101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,272  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/404,950  
FILING DATE: 13-MAR-1995  
APPLICATION NUMBER: 08/336,257  
FILING DATE: 7-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/314,083  
FILING DATE: 28-SEPT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/311,363  
FILING DATE: 23-SEPT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290,012  
FILING DATE: 11-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/223,305  
FILING DATE: 4-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/193,078  
FILING DATE: 07-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,097  
FILING DATE: 5-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/105,536  
FILING DATE: 11-AUG-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,231  
FILING DATE: 13-JULY-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: 10-APR-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06903  
FILING DATE: 14-AUG-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/745,206  
FILING DATE: 15-AUG-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-450-272-8  
Alignment Scores:  
Pred. No.: 14.1  
Score: 81.00  
Percent Similarity: 33.79%  
Best Local Similarity: 21.46%  
Query Match: 7.77%  
Gaps: 4  
Length: 7175  
Matches: 47  
Conservative: 27  
Mismatch: 75  
Indels: 70  
Gaps: 9  
US-09-782-953-7 (1-198) x US-08-450-272-8 (1-7175)  
QY 7 SerTyraAapPheSerLeuIleAlaCysValAlaAaAapPheSerGluSer 26  
Db 2232 AACCTACACTGCTGCTGATGATCTTTCTGCGCATCGCTGTGAGCAACCTGCGCAACGCCCA 2291  
QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyraAapLysAapThrPheGln 46  
Db 2292 GAGCTGACCAAGATGAAGAGAGATGAGAAAGCAAGCAATCAAGACTGCTGCA 2351  
QY 47 TyrPheLysSerPheLysArgValArgIleAaPheSerAaPheLysSerAlaAaAap 66  
Db 2352 AAGGCCAAGAGAGTGGTGAAGTC-----ACGCCATGCTGCGCCGGAAC 2396  
QY 66 ----- 66  
Db 2397 ATCTCCATGCGCGGAGGAGAGCAACTGCGCAAGCGCGCTGCTGAGAGAGCGG 2456  
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81  
Db 2457 GCCAGCACTAGGGGTGAGAACTCGCGGCGCACTGCGAGCCCTGTACAGCAAGT 2516  
QY 82 -----LysLeuTyraPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97

```
Db 2517 GACCCGAGGAGCGCTGGCTTCCCACTAGCGCCACCTG----- 2558
QY 98 ProProAsnProAspGlyGlnPheLeuLeuSerProPro----- 110
Db 2559 -----CGCCCGACATGACAGACGACCTGGACCGCGCTGGTGGACTGGCGCG 2612
QY 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GAGCGCGCGCGGCGCGCTGGAGGCAAGCCCGACCTGAGGCTGGCGAG----- 2663
QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGCGCGCTGCACCTCCGCGAGCGACCGCACCGC 2708
QY 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2709 GACAAGGCAAGACCCCGCGCG-----GGGACACGAGCGCGAGCA 2750
QY 167 -----AsnGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCCCGAAGCGGAGCGCGCGCGCTGCCCGGAGGAGCGCGCGCGCG 2807
```

## RESULT 15

```
US-09-268-163-7
; Sequence 7, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipecombe, Diane
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 7177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 146..6856
US-09-268-163-7
```

## Alignment Scores:

Pred. No.:	14.1	Length:	7177
Score:	81.00	Matches:	47
Percent Similarity:	33.79%	Conservative:	27
Best Local Similarity:	21.46%	Mismatches:	75
Query Match:	7.77%	Indels:	70
DB:	4	Gaps:	9

US-09-782-953-7 (1-198) x US-09-268-163-7 (1-7177)

```
QY 7 SerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2234 AACTACACTCTGCTGATGCTCTTCTGGCCATCGCTGTGCACAACCTGGCCAAACCCCAA 2293
QY 27 GluThrArgAlaLysPheGluSerLysPheArgThrTyrAspLysAspThrThrPheGln 46
Db 2294 GAGCTGACCAAGGATCAAGAGGAGATGGAAAGACGACCAATCAGAAAGCTTCTCTGCA 2353
QY 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2354 AAGGCCAAGAGAGTGGCTCAAGTC-----AGCCCATGCTGCCCGCAAC 2398
QY 66 ----- 66
Db 2399 ATCTCCATCGCCCGCAGGAGGAGACTCGGCCAAGCGCGCTCGGTGTGGAGCAGCGG 2458
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
```

```
Db 2459 GCACAGCCAGCTACGGCTGCAGAACCTGGCGGCGCAGCTGCGAGCGCTGTACAGCGAGATG 2518
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2519 GACCCCGAGGAGCGGCTGGCTTGGCCACTACGCGCCACCTG----- 2560
QY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
Db 2561 -----CGCCCGACATGAAGACGCGACCTGGACCGCGCGCTGGTGGAGCTGGCGCGC 2614
QY 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2615 GACCGCGCGCGCGCGCGCTGGAGGCAAGCCCGACCTGAGGCTGGCGAG----- 2665
QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2666 -----GCCCGCGAGGCGCTGCACCTCCGCGCGAGCGCACCCACCGGACCGC 2710
QY 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2711 GACAAGGCAAGACCCCGCGCG-----GGGACACGAGCGCGAGCA 2752
QY 167 -----AsnGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2753 GAGGCCCGAAGCGGAGCGCGCGCGCTGCCCGGAGGAGCGCGCGCGCGCG 2809
```

Search completed: December 14, 2002, 22:30:25  
Job time : 46.0506 secs

**THIS PAGE BLANK (uspro)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:19:20 ; Search time 33.2872 Seconds  
(without alignments)  
13100.560 Million cell updates/sec

Title: US-09-782-953-1  
Perfect score: 599  
Sequence: 1 gaggtcaaggaaactcca.....gaggaggaggaggagagat 599

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA New:

- 1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413.8	69.1	2050	5	US-09-724-676-12756 Sequence 12756, A
2	413.8	69.1	2050	5	US-09-724-676A-12756 Sequence 12756, A
3	413	68.9	1931	5	US-09-724-676-12748 Sequence 12748, A
4	413	68.9	1931	5	US-09-724-676A-12748 Sequence 12748, A
5	413	68.9	2297	5	US-09-724-676-12747 Sequence 12747, A
6	413	68.9	2297	5	US-09-724-676A-12747 Sequence 12747, A
7	404.8	67.6	2348	6	US-10-240-965-60 Sequence 60, Appl
8	349.2	58.3	1893	5	US-09-724-676-12741 Sequence 12741, A
9	349.2	58.3	1893	5	US-09-724-676A-12741 Sequence 12741, A
10	343.8	57.4	1839	5	US-09-724-676-12743 Sequence 12743, A
11	343.8	57.4	1839	5	US-09-724-676A-12743 Sequence 12743, A
12	343.8	57.4	1876	5	US-09-724-676-12745 Sequence 12745, A
13	343.8	57.4	1876	5	US-09-724-676A-12745 Sequence 12745, A
14	340.4	56.8	911	5	US-09-724-676-12749 Sequence 12749, A
15	340.4	56.8	911	5	US-09-724-676A-12749 Sequence 12749, A
16	338.8	56.6	1213	5	US-09-724-676-12740 Sequence 12740, A
17	338.8	56.6	1213	5	US-09-724-676A-12740 Sequence 12740, A
18	338	56.4	1094	5	US-09-724-676-12750 Sequence 12750, A
19	338	56.4	1094	5	US-09-724-676A-12750 Sequence 12750, A
20	274.2	45.8	1056	5	US-09-724-676-12742 Sequence 12742, A
21	274.2	45.8	1056	5	US-09-724-676A-12742 Sequence 12742, A
22	268.8	44.9	1002	5	US-09-724-676-12744 Sequence 12744, A
23	268.8	44.9	1002	5	US-09-724-676A-12744 Sequence 12744, A
24	268.8	44.9	1039	5	US-09-724-676-12746 Sequence 12746, A
25	268.8	44.9	1039	5	US-09-724-676A-12746 Sequence 12746, A
26	264.4	44.1	442	5	US-09-513-999C-1772 Sequence 1772, Ap

27	203.8	34.0	2411	5	US-09-724-676-15329	Sequence 15329, A
28	203.8	34.0	2411	5	US-09-724-676A-15329	Sequence 15329, A
29	197.6	33.0	3159	6	US-10-290-438-1	Sequence 1, Appli
30	181.8	30.4	615	6	US-10-290-438-8	Sequence 8, Appli
C 31	153	25.5	1021	6	US-10-290-438-3	Sequence 3, Appli
C 32	144.6	24.1	446	6	US-10-203-138A-2101	Sequence 2101, Ap
33	143.6	24.0	1577	5	US-09-724-676-12755	Sequence 12755, A
34	143.6	24.0	1577	5	US-09-724-676A-12755	Sequence 12755, A
35	143.6	24.0	1943	5	US-09-724-676-12752	Sequence 12752, A
36	143.6	24.0	1943	5	US-09-724-676A-12752	Sequence 12752, A
C 37	77.4	12.9	123	6	US-10-203-138A-827	Sequence 827, App
C 38	75.6	12.6	486	6	US-10-203-138A-827	Sequence 827, App
39	72.2	12.1	305	5	US-09-724-676-12751	Sequence 12751, A
40	72.2	12.1	305	5	US-09-724-676A-12751	Sequence 12751, A
41	71	11.9	557	5	US-09-724-676-12753	Sequence 12753, A
42	71	11.9	557	5	US-09-724-676A-12753	Sequence 12753, A
C 43	70.6	11.8	85	6	US-10-203-138A-7216	Sequence 7216, Ap
44	68.6	11.5	740	5	US-09-724-676-12754	Sequence 12754, A
45	68.6	11.5	740	5	US-09-724-676A-12754	Sequence 12754, A

ALIGNMENTS

RESULT 1  
US-09-724-676-12756  
; Sequence 12756, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 12756  
; LENGTH: 2050  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc feature  
; LOCATION: (47)..(47)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12756

Query Match 69.1%; Score 413.8; DB 5; Length 2050;  
Best local Similarity 85.1%; Pred. No. 1e-119;  
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

Qy	39	GAGCGAGTGGTTCGTTAAGCGTCGCCCGTGAAGAGCAGAGATGATTTTAGGACCTTTA	98
Db	141	GACTGCGGGTCTGTGACGCTTTCACTGAAGAGCAGAGATGATTTTAGAAGACTTTA	200
Qy	99	GCTACAAATTTTAGCTCCCTCATTTGTCGCAACGATGATGCTTCAGGAAAGTG	158
Db	201	ACTACAGTTTGTAGCTCCCTCATTTGTCGCAACGATGATGATGATCTTCAGGAAAGTG	260
Qy	159	AGACAGGGGCAAAATTTGAATCCCTCTTCAGAACATATGACAGGACACCCCTTCCAGT	218
Db	261	AAACAGGGGCAAAATTTGAGTCCCTCTTTAGGAGGTATGACAGGACATCCCTTTCAGT	320
Qy	219	ATTTTAAGCTTCAACAGTGTCCGATTAACCTTACGAAACCCCTTATCTGCAGCCGATG	278
Db	321	ATTTTAAGCTTCAACAGGATGAGTAATTAACCTTACGAAACCCCTTCTCCGACGAGATG	380
Qy	279	CCAGGCTGCGGCTGCACAGACCCGAGTTCCTGGGGAGGAAATGAAAGTTGTAATTTGCTC	338
Db	381	CCAGGCTCCAGCTGCATAGAGCTGAGTTTCTGGGAAGGAAATGAAAGTTATATTTTGCTC	440

QY 339 AGACTTACACATAGAGAAAGTTGACACCTGGCTCCGCC-AATCCCGACAAAGATTCTCA 397  
 Db 441 AGACCTTACACATAGAGAAAGTTGACACCTGGCTCCGCCAAATTCAGACAGAGATTCTCA 500  
 QY 338 TCTCCCTCCCGGCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA 457  
 Db 501 TCTCCCTCCCGGCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA 560  
 QY 458 TAAATTACGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAAGTATGAATGC 517  
 Db 561 TAACTATGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAAGTATGAATGC 620  
 QY 518 ATGCAGCCGACAGACACCACTCCGAGTGGTGGTCCACGTGTGAGAGTCAAGAGAGA 577  
 Db 621 AGCAGCCGACGACACCACTCCGAGTGGTGGTCCACGTGTGAGAGTCAAGAGAGA 680  
 QY 578 ATGAGAGAGAGAGAGAA 594  
 Db 681 AGGAGAGAGAGAGAGAA 697

## RESULT 2

US-09-724-676A-12756  
 ; Sequence 12756, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12756  
 ; LENGTH: 2050  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (7)..(7)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (47)..(47)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676A-12756

Query Match 69.1%; Score 413.8; DB 5; Length 2050;  
 Best Local Similarity 85.1%; Pred. No. 1e-119;  
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGGAGTCGTTGCTTAAGCTGCGCCCGTGAAGAAAGCAAGATGATTTTAGGACTTTA 98  
 Db 141 GACTGCGTGGGTCTGTAGCGCTTCACTGTAGAGAAAGAGATGATTTAGAACTTTA 200  
 QY 99 GCTACAAATTAGTCCCTGATGCTGTGTGTGGCAAAAGATGATGTTTCAAGCGAAAGTG 158  
 Db 201 ACTACAGTTTAACTCCCTGATGCTGTGTGTGGCAAAAGATGATGTTTCAAGCGAAAGTG 260  
 QY 159 AGACCAAGGCGCAAAATTGAATCCCTCTTCAAGACATATGACAAAGACACCACTTCCAGT 218  
 Db 261 AAACCAAGGCGCAAAATTGAATCCCTCTTCAAGACATATGACAAAGACATCACTTCCAGT 320  
 QY 219 ATTTTAAGCTTCAAAAGTGTCCGATTAACCTTCAAGCAACCTTATCTGACCGATG 278  
 Db 321 ATTTTAAGCTTCAAAAGTGTCCGATTAACCTTCAAGCAACCTTATCTGACCGATG 380  
 QY 279 CCAAGCTCGGTGTCACAAAGACCGAGTCTGTGGGGAAGAAAGTATGATTTTGTCTC 338  
 Db 381 CCAAGCTCGGTGTCACAAAGACCGAGTCTGTGGGGAAGAAAGTATGATTTTGTCTC 440  
 QY 339 AGACTTACACATAGAGAAAGTTGACACCTGGCTCCGCC-AATCCCGACAAAGATTCTCA 397  
 Db 441 AGACCTTACACATAGAGAAAGTTGACACCTGGCTCCGCCAAATTCAGACAGAGATTCTCA 500

QY 398 TCTCCCTCCCGGCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA 457  
 Db 501 TCTCCCTCCCGGCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA 560  
 QY 458 TAAATTACGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAAGTATGAATGC 517  
 Db 561 TAACTATGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAAGTATGAATGC 620  
 QY 518 ATGCAGCCGACAGACACCACTCCGAGTGGTGGTCCACGTGTGAGAGTCAAGAGAGA 577  
 Db 621 AGCAGCCGACGACACCACTCCGAGTGGTGGTCCACGTGTGAGAGTCAAGAGAGA 680  
 QY 578 ATGAGAGAGAGAGAGAA 594  
 Db 681 AGGAGAGAGAGAGAGAA 697

## RESULT 3

US-09-724-676-12748  
 ; Sequence 12748, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12748  
 ; LENGTH: 1931  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (11)..(11)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (18)..(18)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (23)..(23)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676-12748

Query Match 68.9%; Score 413; DB 5; Length 1931;  
 Best Local Similarity 85.4%; Pred. No. 1.8e-119;  
 Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTCGTTAAGCGTGTGCGCCCGTGAAGAAAGCAAGATGATTTTAGGACTTTAGCTA 102  
 Db 26 GCGTGGCTGTAGCGCTTCACTGTAGAGAAAGCAAGATGATTTTAGAACTTTA 85  
 QY 103 CAATTTAGCTCCCTGATGCTGTGTGTGGCAAAAGATGATGTTTCAAGCGAAAGTGAGAC 162  
 Db 86 CAGTTTAACTCCCTGATGCTGTGTGTGGCAAAAGATGATGTTTCAAGCGAAAGTGAGAC 145  
 QY 163 CAGGCGCAAAATTGAATCCCTCTTCAAGACATATGACAAAGACACCACTTCCAGTATTT 222  
 Db 205 CAGGCGCAAAATTGAATCCCTCTTCAAGACATATGACAAAGACATCACTTCCAGTATTT 205  
 QY 223 TAAGACTTCAAAAGTGTCCGATTAACCTTCAAGCAACCTTATCTGACCGATGCGCAG 282  
 Db 206 TAAGACTTCAAAAGTGTCCGATTAACCTTCAAGCAACCTTATCTGACCGATGCGCAG 265  
 QY 283 GTCGCGCTGACAAAGACCGAGTCTGTGGGGAAGAAAGTATGATTTTGTCTCAGAC 342  
 Db 325 GTCGCGCTGACAAAGACCGAGTCTGTGGGGAAGAAAGTATGATTTTGTCTCAGAC 325  
 QY 343 TTTACATAGAGAAAGTTGACACCTGGCTCCGCC-AATCCCGACAAAGATTCTCAATCTC 401

Db 326 CTTACACATAGGAAGTCAACCTCGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTC 385  
QY 402 CCTCCGGGCTCTCTCCGGTTGGTGGAAACAAGTAGAAGATGCCACCCCGCTCATAAA 461  
Db 386 CCTCCGGGCTCTCCGGCAGTGGATGGAAACAAGTAGAAGATGCCACCCCGCTCATAAA 445  
QY 462 TTACGATCTTTATATGTCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGATGC 521  
Db 446 CTATGATCTCTTATATGTCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGATGC 505  
QY 522 AGCGACAGACACCACTCCCAAGTGTGGTGTCCAGCTGTGTCAGAGTGCACCAAGAGATGA 581  
Db 506 AGCGACTGACACCACTCCCAAGTGTGGTGTCCAGCTGTGTCAGAGTGCACCAAGAGATGA 581  
QY 582 GGAGCAAGAGGAA 594  
Db 566 GGAAGAAGAGGAA 578

## RESULT 4

US-09-724-676A-12748  
; Sequence 12748, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12748  
; LENGTH: 1931  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12748

Query Match 68.9%; Score 413; DB 5; Length 1931;  
Best Local Similarity 85.4%; Pred. No. 1.8e-119;  
Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTCTGTTAAGCGTCTGCCCGTGAAGCAAGATGATTTTAGGACTTTAGCTA 102  
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTAAGAAAGCAAGATGATTTTAGAACTTTAACTA 85  
QY 103 CAATTTTAGCTCCCTGATTCGTTGTGCAAAACAGATGATGCTTCAGCGAAGTGAGAC 162  
Db 86 CAGTTTGTAGCTCCCTGATTCGTTGTGCAAAACAGATGATGCTTCAGCGAAGTGAGAC 145  
QY 163 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTT 222  
Db 146 CAGGCGCAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCCTTTTCAGTATTT 205  
QY 223 TAAGAGCTTCAAAACGTCGCGATAAATTCAGAACCCCTTATCTGACGCGATGCCAG 282  
Db 206 TAAGAGCTTCAAAACGTCGCGATAAATTCAGAACCCCTTCTCCGACGAGATGCCAG 265  
QY 283 GCTGCGGCTGCAAGACCCAGTTCTCGGGGAAGGAATGAAGTTGATTTTGTCTCAGAC 342  
Db 266 GCTCCAGCTGCATAGACTGAGTTCTCGGGAAGGAATGAAGTTGATTTTGTCTCAGAC 325  
QY 343 TTACACATAGGAAGTTCACACCTGGCTCCGCC - AATCCGCAACAAACAGTTCTCTCATCTC 401

Db 326 CTTACACATAGGAAGTCAACCTCGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTC 385  
QY 402 CCTCCGGGCTCTCTCCGGTTGGTGGAAACAAGTAGAAGATGCCACCCCGCTCATAAA 461  
Db 386 CCTCCGGGCTCTCCGGCAGTGGATGGAAACAAGTAGAAGATGCCACCCCGCTCATAAA 445  
QY 462 TTACGATCTTTATATGTCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGATGC 521  
Db 446 CTATGATCTCTTATATGTCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGCGATGC 505  
QY 522 AGCGACAGACACCACTCCCAAGTGTGGTGTCCAGCTGTGTCAGAGTGCACCAAGAGATGA 581  
Db 506 AGCGACTGACACCACTCCCAAGTGTGGTGTCCAGCTGTGTCAGAGTGCACCAAGAGATGA 581  
QY 582 GGAGCAAGAGGAA 594  
Db 566 GGAAGAAGAGGAA 578

## RESULT 5

US-09-724-676-12747  
; Sequence 12747, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12747  
; LENGTH: 2297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12747

Query Match 68.9%; Score 413; DB 5; Length 2297;  
Best Local Similarity 85.4%; Pred. No. 1.9e-119;  
Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTCTGTTAAGCGTCTGCCCGTGAAGCAAGATGATTTTAGGACTTTAGCTA 102  
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTAAGAAAGCAAGATGATTTTAGAACTTTAACTA 85  
QY 103 CAATTTTAGCTCCCTGATTCGTTGTGCAAAACAGATGATGCTTCAGCGAAGTGAGAC 162  
Db 86 CAGTTTGTAGCTCCCTGATTCGTTGTGCAAAACAGATGATGCTTCAGCGAAGTGAGAC 145  
QY 163 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTT 222  
Db 146 CAGGCGCAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCCTTTTCAGTATTT 205  
QY 223 TAAGAGCTTCAAAACGTCGCGATAAATTCAGAACCCCTTATCTGACGCGATGCCAG 282  
Db 206 TAAGAGCTTCAAAACGTCGCGATAAATTCAGAACCCCTTCTCCGACGAGATGCCAG 265  
QY 283 GCTGCGGCTGCAAGACCCAGTTCTCGGGGAAGGAATGAAGTTGATTTTGTCTCAGAC 342  
Db 266 GCTCCAGCTGCATAGACTGAGTTCTCGGGAAGGAATGAAGTTGATTTTGTCTCAGAC 325

QY 343 TTACACATAGAGAGTTGACACCTGCTCCGCC-AATCCCGACAAACAGTTCTCATCTC 401  
DB 326 CTACACATAGAGAGTTGACACCTGCTCCGCCAAATCCAGACAAAGAGTTCTCATCTC 385  
QY 402 CCTCCCGGCTCTCTCCGCTGAGGAAACAGTAGAAGATGCCACCCCGCTCATATAA 461  
DB 386 CCTCCCGGCTCTCTCCGCTGAGGAAACAGTAGAAGATGCCACCCCGCTCATATAA 445  
QY 462 TTACGATCTTTATATGTCATCTCCAAAGTGGGGCCAGAGAGATGATGATGATG 521  
DB 446 CTATGATCTCTTATATGTCATCTCCAAAGTGGGGCCAGAGAGATGATGATGATG 505  
QY 522 AGCGACAGACACCACTCCGAGTGTGTGTCCAGCTGTGTGAGATGACCAAGAGATGA 581  
DB 506 AGCGACAGACACCACTCCGAGTGTGTGTCCAGCTGTGTGAGATGATGATGATGA 565  
QY 582 GGAGAGAGAGGAA 594  
DB 566 GGAGAGAGAGGAA 578

## RESULT 6

US-09-724-676A-12747

Sequence 12747, Application US/09724676A

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676A

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12747

LENGTH: 2297

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (11)..(11)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (18)..(18)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (23)..(23)

OTHER INFORMATION: n is a,c,g, or t

US-09-724-676A-12747

Query Match

Best Local Similarity 68.9%; Score 413; DB 5; Length 2297;

Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTGTTAAGCGTCTGCCCCGTAAGAAAGCAATATTTTGGGACTTTACTTA 102  
DB 26 GCGTGGGCTCTAGCGCTTCACTGTAGAAAGCAAGTGCATTTTAACTTA 85  
QY 103 CAATTTAGCTCCCTGATGCTGTGTGTGGCAAGCATGATGCTTTACGAGAAAGTGAGAC 162  
DB 86 CAGTTTAGCTCCCTGATGCTGTGTGTGGCAAGCATGATGCTTTACGAGAAAGTGAGAC 145  
QY 163 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCAGTATTT 222  
DB 146 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCAGTATTT 205  
QY 223 TAAAGAGCTTCAAAAGTGTCCGATTAAGTTCAGCAACCCCTTATGAGCGAGATGCGAG 282  
DB 206 TAAAGAGCTTCAAAAGTGTCCGATTAAGTTCAGCAACCCCTTATGAGCGAGATGCGAG 265  
QY 283 GGTGGGCTGCGACAAAGACGAGTTCCTGGGGAAGAAATGAATGTTATTTTGTCTCAGAC 342  
DB 266 GGTGGGCTGCGACAAAGACGAGTTCCTGGGGAAGAAATGAATGTTATTTTGTCTCAGAC 325

QY 343 TTACACATAGAGAGTTGACACCTGCTCCGCC-AATCCCGACAAACAGTTCTCATCTC 401  
DB 326 CTACACATAGAGAGTTGACACCTGCTCCGCCAAATCCAGACAAAGAGTTCTCATCTC 385  
QY 402 CCTCCCGGCTCTCTCCGCTGAGGAAACAGTAGAAGATGCCACCCCGCTCATATAA 461  
DB 386 CCTCCCGGCTCTCTCCGCTGAGGAAACAGTAGAAGATGCCACCCCGCTCATATAA 445  
QY 462 TTACGATCTTTATATGTCATCTCCAAAGTGGGGCCAGAGAGATGATGATGATG 521  
DB 446 CTATGATCTCTTATATGTCATCTCCAAAGTGGGGCCAGAGAGATGATGATGATG 505  
QY 522 AGCGACAGACACCACTCCGAGTGTGTGTCCAGCTGTGTGAGATGACCAAGAGATGA 581  
DB 506 AGCGACAGACACCACTCCGAGTGTGTGTCCAGCTGTGTGAGATGATGATGATGA 565  
QY 582 GGAGAGAGAGGAA 594  
DB 566 GGAGAGAGAGGAA 578

## RESULT 7

US-10-240-965-60

Sequence 60, Application US/10240965

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: SHIFMAN, Dov

APPLICANT: SOMOGYI, Roland

APPLICANT: LAWN, Richard M.

APPLICANT: SEITHAMER, Jeffrey J.

APPLICANT: PORTER, Gordon J.

APPLICANT: MIKITA, Thomas

TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION

FILE REFERENCE: PA-0025 PCT

CURRENT APPLICATION NUMBER: US/10/240,965

PRIOR FILING DATE: 2002-10-04

NUMBER OF SEQ ID NOS: 276

SOFTWARE: PERL Program

SEQ ID NO 60

LENGTH: 2348

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 042176.5

US-10-240-965-60

Query Match

Best Local Similarity 67.6%; Score 404.8; DB 6; Length 2348;

Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

QY 39 GAGCGAGTCGTTGTTAAGCGTCTGCCCCGTAAGAAAGCAATATTTTGGGACTTTTA 98  
DB 110 GAGTCGTTGTTAAGCGTCTGCCCCGTAAGAAAGCAATATTTTGGGACTTTTA 169  
QY 99 GCTACATTTTAACTCCCTGATGCTGTGTGTGGCAAGCATGATGCTTTACGAGAAAGTG 158  
DB 170 ACTACATTTTAACTCCCTGATGCTGTGTGTGGCAAGCATGATGCTTTACGAGAAAGTG 229  
QY 159 AGACCGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCAGT 218  
DB 230 AGACCGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCAGT 289  
QY 219 ATTTAAGAGCTTCAAAAGTGTCCGATTAAGTTCAGCAACCCCTTATGAGCGAGATG 278  
DB 290 ATTTAAGAGCTTCAAAAGTGTCCGATTAAGTTCAGCAACCCCTTATGAGCGAGATG 349  
QY 279 CGAGGCTGCGCTGCGACAAAGACGAGTTCCTGGGGAAGAAATGAATGTTATTTTGTCTC 338





CURRENT APPLICATION NUMBER: US/09/724.676  
 : CURRENT FILING DATE: 2000-11-28  
 : NUMBER OF SEQ ID NOS: 97222  
 : SOFTWARE: PatentIn version 3.2  
 : SEQ ID NO 12743  
 : LENGTH: 1839  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-724-676-12743

Query Match 57.4%; Score 343.8; DB 5; Length 1839;  
 Best Local Similarity 88.9%; Pred. No. 9.1e-98;  
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCCCTTCAGATATTTTA 224  
 Db 56 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCCCTTCAGATATTTTA 115  
 QY 225 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTATCTGACGCCGATGCCAGGC 284  
 Db 116 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTATCTGACGCCGATGCCAGGC 175  
 QY 285 TGGGCTGCAACAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344  
 Db 176 TCCAGCTGATTAAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 235  
 QY 345 TACACATAGGAAGTTTCAACCTTGGCTCCGCC-ATCCCGCAACAAGCTTCATCTCCC 403  
 Db 236 TACACATAGGAAGTTTCAACCTTGGCTCCGCCAATCCAGACAGATTTCTGATCTCCC 295  
 QY 404 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 463  
 Db 296 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 355  
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGATCTGCATGCCAG 523  
 Db 356 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGATCTGCATGCCAG 415  
 QY 524 CGACAGACCACTCCGAGTGGTGGTCCAGCGTGTGAGTGCACCAAGAAATGAG 583  
 Db 416 CGACAGACCACTCCGAGTGGTGGTCCAGCGTGTGAGTGCACCAAGAAATGAG 475  
 QY 584 AGAAGAGGAA 594  
 Db 476 AAGAAGAGGAA 486

RESULT 11  
 : US-09-724-676A-12743  
 : Sequence 12743, Application US/09724676A  
 : GENERAL INFORMATION:  
 : APPLICANT: Compugen LTD  
 : TITLE OF INVENTION: Variance of alternative splicing  
 : FILE REFERENCE: 129181.4 Compugen  
 : CURRENT APPLICATION NUMBER: US/09/724.676A  
 : CURRENT FILING DATE: 2000-11-28  
 : NUMBER OF SEQ ID NOS: 97222  
 : SOFTWARE: PatentIn version 3.2  
 : SEQ ID NO 12743  
 : LENGTH: 1839  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-724-676A-12743

Query Match 57.4%; Score 343.8; DB 5; Length 1839;  
 Best Local Similarity 88.9%; Pred. No. 9.1e-98;  
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;  
 QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCACTTCAGATATTTTA 224  
 Db 56 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCACTTCAGATATTTTA 115  
 QY 225 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTATCTGACGCCGATGCCAGGC 284

Db 116 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTTCAGAGACGATGCCAGGC 175  
 QY 285 TGGGCTGCAACAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344  
 Db 176 TCCAGCTGATTAAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 235  
 QY 345 TACACATAGGAAGTTTCAACCTTGGCTCCGCC-ATCCCGCAACAAGCTTCATCTCCC 403  
 Db 236 TACACATAGGAAGTTTCAACCTTGGCTCCGCCAATCCAGACAGATTTCTGATCTCCC 295  
 QY 404 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 463  
 Db 296 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 355  
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGAACTGATGCCAG 523  
 Db 356 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGAACTGATGCCAG 415  
 QY 524 CGACAGACCACTCCGAGTGGTGGTCCAGCTGTGTGAGTGCACCAAGAAATGAG 583  
 Db 416 CGACAGACCACTCCGAGTGGTGGTCCAGCTGTGTGAGTGCACCAAGAAATGAG 475  
 QY 584 AGAAGAGGAA 594  
 Db 476 AAGAAGAGGAA 486

RESULT 12  
 : US-09-724-676-12745  
 : Sequence 12745, Application US/09724676  
 : GENERAL INFORMATION:  
 : APPLICANT: Compugen LTD  
 : TITLE OF INVENTION: Variance of alternative splicing  
 : FILE REFERENCE: 129181.4 Compugen  
 : CURRENT APPLICATION NUMBER: US/09/724.676  
 : CURRENT FILING DATE: 2000-11-28  
 : NUMBER OF SEQ ID NOS: 97222  
 : SOFTWARE: PatentIn version 3.2  
 : SEQ ID NO 12745  
 : LENGTH: 1876  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-724-676-12745

Query Match 57.4%; Score 343.8; DB 5; Length 1876;  
 Best Local Similarity 88.9%; Pred. No. 9.2e-98;  
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCACTTCAGATATTTTA 224  
 Db 93 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGACACCACTTCAGATATTTTA 152  
 QY 225 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTATCTGACGCCGATGCCAGGC 284  
 Db 153 AAGAGCTTCAAGCTGTCGGGATTAACCTTACGCAACCCCTTATCTGACGCCGATGCCAGGC 212  
 QY 285 TGGGCTGCAACAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344  
 Db 213 TCCAGCTGATTAAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 272  
 QY 345 TACACATAGGAAGTTTCAACCTTGGCTCCGCC-ATCCCGCAACAAGCTTCATCTCCC 403  
 Db 273 TACACATAGGAAGTTTCAACCTTGGCTCCGCCAATCCAGACAGATTTCTGATCTCCC 332  
 QY 404 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 463  
 Db 333 CTCGGGCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGTCATTAATT 392  
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGAACTGATGCCAGGC 523  
 Db 393 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGATGATGAACTGATGCCAGGC 452

QY 524 CGACAGACACCACTCCCAAGTGTGGGTCCAGTGTGTGAGAGTGAACCAAGAGAAATGAGG 583  
Db 453 CGACTGACACCACTCCCAAGTGTGGGTCCATGTATGTGAGAGTGAATCAAGAGAGGAGG 512  
QY 584 AGGAAGAGGAA 594  
Db 513 AAGAAGAGGAA 523

## RESULT 13

US-09-724-676A-12745  
; Sequence 12745, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12745  
; LENGTH: 1876  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676A-12745

Query Match 57.4%; Score 343.8; DB 5; Length 1876;  
Best Local Similarity 88.9%; Pred. No. 9.2e-98;  
Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;  
QY 165 GGGCCAAATTTGAATCCCTCTTCAGACATATGACAGGACACCACTTCAGTATTTTA 224  
Db 93 GGGCCAAATTTGATCCCTCTTTAGGACGTATGACAGGACATCACCTTCAGTATTTTA 152  
QY 225 AGAGCTTCAAAACGTGTCGGGATAAACTTCAGCAACCCCTTATGTCAGCGCGATGCCAGGC 284  
Db 153 AGAGCTTCAAAACGTGTCAGATAAACTTCAGCAACCCCTTCTCGCAGAGATGCCAGGC 212  
QY 285 TGGCGTGCACAGACCGAGTCTCTGGGAGGAATGAAGTTGTATTTTGTCTCAGACTT 344  
Db 213 TCCAGCTGCAATAGACTGAGTTCTTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACTT 272  
QY 345 TACACATAGGAAGTTACACTGTGCTCCGCC-AATCCGCAACAGTTCCTCATCTCCC 403  
Db 273 TACACATAGGAAGTTACACTGTGCTCCGCCAAATCCAGACAGAGTTCCTCATCTCCC 332  
QY 404 CTCGGGCTCTCTCCCGTGTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAAT 463  
Db 333 CTCGGGCTCTCCCGAGTGGGATGGAACAAGTGAAGATGCCACCCCGTCAATAACT 392  
QY 464 ACAGTCTTTATATGCAATCTCCAAGTGGGCGCAGAGAGAGTGAATGCACTGCAG 523  
Db 393 ATGATCTTTATATGCAATCTCCAAGTGGGCGCAGAGAGAGTGAATGCACTGCAG 452  
QY 524 CGACAGACCACTCCCAAGTGTGGTGTCCACGTGTGAGAGTGAACCAAGAGAGTGAAG 583  
Db 453 CGACTGACACCACTCCCAAGTGTGGTGTCCACGTGTGAGAGTGAACCAAGAGAGTGAAG 512  
QY 584 AGGAAGAGGAA 594  
Db 513 AAGAAGAGGAA 523

## RESULT 14

US-09-724-676-12749  
; Sequence 12749, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12749  
; LENGTH: 911  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11)-(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)-(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (23)-(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12749

Query Match 56.8%; Score 340.4; DB 5; Length 911;  
Best Local Similarity 84.5%; Pred. No. 8e-97; Mismatches 0; Indels 1; Gaps 1;  
Matches 394; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 43 GAGTCGTTTAAAGCGTCTGCCCGTGAAAGAGCAGAAATGATTTAGGCACTTTAGCTA 102  
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTAAAGAAAGCAAGATGCAATTTAGAACTTTAACTA 85  
QY 103 CAATTTAGCTCTCGATTTGTTGTGGCAACAGATGATGTTTCAGCGAAAGTGAGAC 162  
Db 86 CAGTTTATAGCTCTCCCTGATTCCTGTGGCAACAGTATATCTTCAGCGAAAGTGAAAC 145  
QY 163 CAGGGCCAAATTTGAATCCCTCTTCAGACATATGACAGGACACCACTTCAGTATTT 222  
Db 146 CAGGGCCAAATTTGATTCCTCTTTAGGACGTATGACAGGACATCACCTTTTCACTATTT 205  
QY 223 TAAGAGCTTCAAAACGTGTCGGATAAACTTCAGCAACCCCTTATTCGACGCCATGCCAG 282  
Db 206 TAAGAGCTTCAAAACGTGTCAGATAAACTTCAGCAACCCCTTCTCGCAGCAGATGCCAG 265  
QY 283 GCTGGCGTGCACAGACCGAGTCTCTGGGAGGAATGAAGTTGTATTTTGTCTCAGAC 342  
Db 266 GCTCAGCTGCAATAGACTGAGTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGAC 325  
QY 343 TTACACATAGGAAGTTACACTGTGCTCCGCC-AATCCGCAACAGTTCCTCATCTC 401  
Db 326 CTTCACATAGGAAGTCTACACTGTGCTCCGCCAAATCCAGACAGAGTTCCTCATCTC 385  
QY 402 CCTCCGGCTCTCTCCCGTGTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAA 461  
Db 386 CCTCCCGCTCTCCCGCAGTGGGATGGAACAAGTGAAGATGCCACCCCGTCAATAA 445  
QY 462 TTACGATCTTTATATGCAATCTCCAAGTGGGCGCAGAGAGAG 507  
Db 446 CTATGATCTTTATATGCAATCTCCAAGTGGGCGCAGAGAGAG 491

## RESULT 15

US-09-724-676A-12749  
; Sequence 12749, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12749  
; LENGTH: 911  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature

```
LOCATION: (11)-(11)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (18)-(18)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)-(23)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12749
```

```
Query March 56.8%; Score 340.4; DB 5; Length 911;
Best Local Similarity 84.5%; Pred. No. 8e-97;
Matches 394; Conservative 0; Mismatches 71; Indels 1; Gaps 1;
```

```
OY 43 GAGTCGTTGTTAGCGCTGCTGCCCCGCGTGAAGAAAGCAAGATGATTTTAGGACTTTAGCTA 102
DB 26 GGTGGGCTGTGAGCGCTTCACTGAAGAAAGCAAGATGATTTAGAACTTTAAGCTA 85
OY 103 CAATTTAGCTCCCTGATTCCTTGTGTGCGCAACGATGATGCTTCAAGGAAAGTGAAC 162
DB 86 CAGTTTAGCTCCCTGATTCCTTGTGTGCGCAACGATGATGCTTCAAGGAAAGTGAAC 145
OY 163 CAGGCGCAATTTGAATCCCTCTCAGAACATATGACAGACACCCCTTCAGATTT 222
DB 146 CAGGCGCAATTTGAATCCCTCTCAGAACATATGACAGACACCCCTTCAGATTT 205
OY 223 TAAGAGCTTCAACGCTGTCGGATTAACCTTCAGCAACCCCTTATCTGACGCCGATGCCAG 282
DB 206 TAAGAGCTTCAACGCTGTCAGATTAACCTTCAGCAACCCCTTCTCGCAGCAGATGCCAG 265
OY 283 GGTGGGCTGCAACAGACCGAGTTCCTGCGGAAGAAATGAAGTTGATTTGCTCAGAC 342
DB 266 GGTCCAGCTGCATGAGACTGAGTTTCTGGAAAGGAAATGAAGTTATTTGCTCAGAC 325
OY 343 TTACACATAGGAAGTTCAACCTGCTCGGC-AAATCCGACAAACAGTTCTCATCTC 401
DB 326 CTTACACATAGGAAGTTCAACCTGCTCGGC-AAATCCGACAAACAGTTCTCATCTC 385
OY 402 CCCTCCGCTCTCTCTCCCTGCTGCGAAACAAAGTAGAAGATGCCACCCCGCTATAA 461
DB 386 CCCTCCGCTCTCTCTCCCTGCTGCGAAACAAAGTAGAAGATGCCACCCCGCTATAA 445
OY 462 TTACGATCTTTATATGATCCATCCAAAGCTGCGGCGCAGAGAGAGAG 507
DB 446 CTTATGATCTTTATATGATCCATCCAAAGCTGCGGCGCAGAGAGAGAG 491
```

Search completed: December 14, 2002, 19:59:16  
Job time : 37.2872 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 33.0556 Seconds  
(without alignments)  
2368.771 Million cell updates/sec

Title: US-09-782-953-7  
Perfect score: 1043  
Sequence: 1 MDRFDSYFSSLIACVAND.....RPKPKIQRPEYTPHLS 198

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DRV=xlh  
-Q=/cgn2/USPTO.spool/US09782953/runat\_11122002\_114431\_17465/app\_query.fasta\_1.1173  
-DB=Published Applications NA -QFWT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09782953 -CGCN 1 1 38 @runat\_11122002\_114431\_17465  
-NCFU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

```
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	597	10	US-09-782-953-5
2	970.5	93.0	2331	10	Sequence 5, Appli
3	970.5	93.0	2355	12	Sequence 11, Appl
4	920.5	88.3	597	10	Sequence 255, App
					Sequence 2, Appli

5	875	83.9	2212	10	US-09-782-953-14	Sequence 14, Appli
6	868	83.2	599	10	US-09-782-953-1	Sequence 1, Appli
7	847.5	81.3	2173	10	US-09-880-107-3340	Sequence 3340, Ap
8	830	79.6	2358	10	US-09-925-302-347	Sequence 347, App
9	632.5	60.6	594	10	US-09-782-953-8	Sequence 8, Appli
10	615.5	59.0	3184	10	US-09-954-456-497	Sequence 497, App
11	615.5	59.0	3184	10	US-09-782-953-17	Sequence 17, Appl
12	592.5	56.8	828	10	US-09-782-953-20	Sequence 20, Appl
13	591.5	56.7	720	10	US-09-782-953-23	Sequence 23, Appl
14	285.5	27.4	412	10	US-09-864-761-10388	Sequence 10388, A
15	285.5	27.4	446	10	US-09-864-761-2064	Sequence 2064, Ap
16	266.5	25.6	486	10	US-09-864-761-809	Sequence 809, App
17	200	19.2	365	10	US-09-728-445-736	Sequence 736, App
18	187.5	18.0	123	10	US-09-864-761-17592	Sequence 17592, A
19	181	17.4	111	10	US-09-864-761-27019	Sequence 27019, A
20	140	13.4	85	10	US-09-864-761-18808	Sequence 18808, A
21	90.5	8.7	12308	9	US-09-854-133-422	Sequence 422, App
22	90.5	8.7	12308	10	US-09-738-973-422	Sequence 422, App
23	87	8.3	4600	9	US-09-736-457-1797	Sequence 1797, Ap
24	87	8.3	4600	9	US-09-902-941-1797	Sequence 1797, Ap
25	82.5	7.9	530	10	US-09-864-761-15897	Sequence 15897, A
26	81.5	7.8	2095	12	US-10-044-090-665	Sequence 665, App
27	81	7.8	3138	10	US-09-880-107-1716	Sequence 1716, App
28	81	7.8	7177	12	US-10-033-026-7	Sequence 7, Appli
29	81	7.8	7364	10	US-09-954-456-1179	Sequence 1179, Ap
30	81	7.8	7364	12	US-10-033-026-5	Sequence 5, Appli
31	81	7.8	7376	12	US-10-033-026-3	Sequence 3, Appli
32	81	7.8	640681	10	US-09-790-988-1	Sequence 1, Appli
33	80.5	7.7	6799	9	US-09-902-941-1883	Sequence 1883, Ap
34	80	7.7	552	10	US-09-954-456-594	Sequence 594, App
35	79	7.6	3775	10	US-09-954-456-2210	Sequence 2210, Ap
36	79	7.6	3775	10	US-09-880-107-2221	Sequence 2221, Ap
37	79	7.6	36651	10	US-09-964-469-3	Sequence 3, Appli
38	78.5	7.5	569	9	US-10-040-739-301	Sequence 301, App
39	77.5	7.4	4108	10	US-09-883-096-1	Sequence 1, Appli
40	77.5	7.4	32191	10	US-09-764-877-3374	Sequence 3374, Ap
41	77	7.4	1440	9	US-09-712-363-94	Sequence 94, Appl
42	77	7.4	2149	10	US-09-880-107-3031	Sequence 3031, Ap
43	77	7.4	5739	10	US-09-960-253-142	Sequence 142, App
44	77	7.4	6417	10	US-09-962-436-288	Sequence 288, App
45	76.5	7.3	780	9	US-09-938-842A-1412	Sequence 1412, Ap

#### ALIGNMENTS

RESULT 1  
US-09-782-953-5  
; Sequence 5, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR FILING DATE: 60/216,601  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(594)  
US-09-782-953-5

Alignment Scores:  
Pred. No.: 1,33e-118 Length: 597  
Score: 1043.00 Matches: 198

Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 10  
 Gaps: 0

US-09-782-953-7 (1-198) x US-09-782-953-5 (1-597)

QY 1 MetAaPheAaPheSerTyraPheSerSerLeuLeuAaCyAaValAaAaAaP 20  
 DB 1 ATGATTTTGGGACTTTTACCTAATTTTACCTCCGATTTGCTGTGGCAACAT 60  
 QY 21 AaPValPheSerGluSerGluThrAaValAaPheGluSerLeuPheArgThrTyraP 40  
 DB 61 GATGCTTCAGCGAAAGTGAAGACAGGGCCAAATTTGAATCCCTCTTCAAGACATATGAC 120  
 QY 41 LysAaPThrThrPheGlnTyraPheLysSerPheLysValArgIlaAaPheSerAa 60  
 DB 121 AAGGACACACACCTTCAGATTTTAAAGACTTCAACGTCGGATTAACCTTCAGCAAC 180  
 QY 61 ProLeuSerAlaAaPValAaArgLeuAaGluAaValAaGluPheLeuGluLysGlu 80  
 DB 181 CCTTATCTGACAGCCATGCGAGCTGCGGCTGCAACAGAGTTCCTGGGAAGGAA 240  
 QY 81 MetLysLeuTyraPheAaGlnThrLeuAaValAaGlySerSerHisLeuAaProProA 100  
 DB 241 ATGAACTTGTATTTGCTCAGACTTACACATAGAAAGTTTCACTGCTCCGCCAAT 300  
 QY 101 ProAaPValGlnPheLeuLeuSerProAaLeuSerProAaValGluTyraPValAa 120  
 DB 301 CCGGACAAACAGTCTCACTCTCCCTCGGCTCTCTCTCCCTGGCTGGAAACAAGTA 360  
 QY 121 GluAaPValThrProValIlaAaPValAaPheLeuLeuTyraAaIleSerLysLeuGlyPro 140  
 DB 361 GAGAGTCCACCCCGCTCAATTAATGATCTTTATATGCACTTCCAAAGCTGGGAGCA 420  
 QY 141 GlyGluLysTyraGluLeuAaValAaAaThrAaPProThrProSerAaValAaHisVal 160  
 DB 421 GGAGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 QY 161 CysGluSerAaPValGlnAaGluAaGluAaGluAaGluAaGluAaGluAaGluAaGlu 180  
 DB 481 TGTGAGAGTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
 QY 181 LysProLysIleIleGlnThrAaArgProGluTyraThrProIleHisLeuSer 198  
 DB 541 AAGCCAAATATCCAGACAGGAGCCGAGTACACACCGATCCACCTTAC 594

## RESULT 2

US-09-782-953-11  
 Sequence 11, Application US/09782953  
 Patent No. US20020150953A1  
 GENERAL INFORMATION:  
 APPLICANT: WILLIAMS, R. SANDERS  
 APPLICANT: ROTHERMEL, BEVERLY  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 FILE REFERENCE: US/09-782-953-11  
 CURRENT APPLICATION NUMBER: US/09-782-953  
 PRIOR FILING DATE: 2001-02-13  
 PRIOR APPLICATION NUMBER: 60/216,601  
 NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 11  
 LENGTH: 2331  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (144) ... (734)  
 US-09-782-953-11

Alignment Scores:

Pred. No.: 7.13e-109  
 Score: 970.50  
 Percent Similarity: 96.97%  
 Best Local Similarity: 93.43%  
 Query Match: 93.05%  
 DB: 10  
 Gaps: 1

US-09-782-953-7 (1-198) x US-09-782-953-11 (1-2331)

QY 1 MetAaPheAaPheSerTyraPheSerSerLeuLeuAaCyAaValAaAaAaP 20  
 DB 144 ATGATTTTGGGACTTTTACCTAATTTTACCTCCGATTTGCTGTGGCAACAT 203  
 QY 21 AaPValPheSerGluSerGluThrAaValAaPheGluSerLeuPheArgThrTyraP 40  
 DB 204 GATGCTTCAGCGAAAGTGAAGACAGGGCCAAATTTGAATCCCTCTTCAAGACATATGAC 263  
 QY 41 LysAaPThrThrPheGlnTyraPheLysSerPheLysValArgIlaAaPheSerAa 60  
 DB 264 AAGGACACACACCTTCAGATTTTAAAGACTTCAACGTCGGATTAACCTTCAGCAAC 323  
 QY 61 ProLeuSerAlaAaPValAaArgLeuAaGluAaValAaGluPheLeuGluLysGlu 80  
 DB 324 CCTTATCTGACAGCCATGCGAGCTGCGGCTGCAACAGAGTTCCTGGGAAGGAA 383  
 QY 81 MetLysLeuTyraPheAaGlnThrLeuAaValAaGlySerSerHisLeuAaProProA 100  
 DB 384 ATGAACTTGTATTTGCTCAGACTTACACATAGAAAGTTTCACTGCTCCGCCAAT 443  
 QY 101 ProAaPValGlnPheLeuLeuSerProAaLeuSerProAaValGluTyraPValAa 120  
 DB 444 CCGGACAAACAGTCTCACTCTCCCTCGGCTCTCTCTCCCTGGCTGGAAACAAGTA 503  
 QY 121 GluAaPValThrProValIlaAaPValAaPheLeuLeuTyraAaIleSerLysLeuGlyPro 140  
 DB 504 GAGAGTCCACCCCGCTCAATTAATGATCTTTATATGCACTTCCAAAGCTGGGAGCA 563  
 QY 141 GlyGluLysTyraGluLeuAaValAaAaThrAaPProThrProSerAaValAaHisVal 160  
 DB 564 GGGGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 623  
 QY 161 CysGluSerAaPValGlnAaGluAaGluAaGluAaGluAaGluAaGluAaGluAaGlu 180  
 DB 624 TGTGAGAGTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680  
 QY 181 LysProLysIleIleGlnThrAaArgProGluTyraThrProIleHisLeuSer 198  
 DB 681 AAGCCAAATATCCAGACAGGAGCCGAGTACACACCGATCCACCTTAC 734

## RESULT 3

US-10-044-090-255  
 Sequence 255, Application US/10044090  
 Patent No. US20020137081A1  
 GENERAL INFORMATION:  
 APPLICANT: Olga Bandman  
 TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
 FILE REFERENCE: PA-0028 US  
 CURRENT APPLICATION NUMBER: US/10-044,090  
 CURRENT FILING DATE: 2002-01-09  
 NUMBER OF SEQ ID NOS: 850  
 SOFTWARE: PERL Program  
 SEQ ID NO 255  
 LENGTH: 2355  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889C81  
 US-10-044-090-255

Alignment Scores:  
 Pred. No.: 7.24e-109  
 Score: 970.50  
 Length: 2355  
 Matches: 185

Percent Similarity: 96.97%  
 Best Local Similarity: 93.43%  
 Query Match: 93.05%  
 DB: 12  
 Indels: 1  
 Gaps: 1

US-09-782-953-7 (1-198) x US-10-044-090-255 (1-2355)  
 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20  
 151 ATGCATTTTGAACATTTTAACTACAGTTTATAGCTCCCTGATTGCTGTGTGCAACAGT 210  
 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 211 GATATCTTACGGAAGTGAACCAAGGCCCAATTTGAGTCCCTCTTTAGGACGATGAC 270  
 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 271 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGANTAACTTCAGCAAC 330  
 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 331 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAA 390  
 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100  
 391 ATGAAGTTATATTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAT 450  
 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120  
 451 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAGTG 510  
 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
 511 GAAGATCGAGCCAGTCATAAATATGATCTCTTATATGCCATCTCCAAAGCTGGGCGCA 570  
 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160  
 571 GGGGAAAAGTATGAATTCAGCAGCAGCAGTGCACCACTCCAGCGTGGTGGTCCATGTA 630  
 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180  
 631 TGTGAGAGTATCAAGAG---AAGGAGGAAGAGAGGAAATGGAAGATGAGGAGACCT 687  
 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
 688 AAGCCAAAATTTATCCAGACAGAGCGCGAGTACACGCCGATCCACCTCAGC 741

## RESULT 4

US-09-782-953-2  
 ; Sequence 2, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; FILE REFERENCE: US/09782953  
 ; CURRENT APPLICATION NUMBER: US/09782953  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 597  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(594)  
 US-09-782-953-2

## Alignment Scores:

Pred. No.: 1-23e-103  
 Score: 920.50  
 Percent Similarity: 92.42%  
 Best Local Similarity: 89.90%  
 Query Match: 88.26%  
 DB: 10  
 Indels: 3  
 Gaps: 1

US-09-782-953-7 (1-198) x US-09-782-953-2 (1-597)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20  
 Db 10 GTGGATCTCGAGGACCTGCCG-----AGCGCCACCATCGCTGCCACCTGGACCCG 60  
 Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 Db 61 CGCGTGTTCGTGCGCGCTGCGCGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120  
 Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 Db 121 AAGGACACCATCTCCAGTATTTTAAAGAGCTTCAACCGTGTCCGGATAAATTCAGCAAC 180  
 Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 Db 181 CCTTATCTGCGAGCGATGCCAGGCTGCGGCTGCACAAAGACCGAGTTCTCGGGAAGGAA 240  
 Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100  
 Db 241 ATGAAGTTGATTTTGTCTCAGACTTTACATAGGAAGTTTACACCTGGCTCCGCCAAT 300  
 Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120  
 Db 301 CCGGACAAACAGTCTCTCATCTCCCTCCGCTCTCTCCGTTGGTGGTGGAAACAGTA 360  
 Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
 Db 361 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATATGCCATCTCCAAAGCTGGGCGCA 420  
 Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160  
 Db 421 GGAGAGAGTATGAATTCAGTGCAGCGACAGCCCACTCCAGCTGGTGGTGGTGGTGGTGG 480  
 Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180  
 Db 481 TGTGAGAGTACCAAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198  
 Db 541 AAGCCAAAATTTATCCAGACAGGCGCGAGTACACCGGATCCACCTTAGC 594

## RESULT 5

US-09-782-953-14  
 ; Sequence 14, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; FILE REFERENCE: US/09782953  
 ; CURRENT APPLICATION NUMBER: US/09782953  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 2212  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (25)..(615)  
 US-09-782-953-14

Alignment Scores:  
 Pred. No.: 3.08e-97 Length: 2212  
 Score: 875.00 Matches: 171  
 Percent Similarity: 90.40% Conservative: 8  
 Best Local Similarity: 86.36% Mismatches: 15  
 Query Match: 83.89% Indels: 4  
 DB: 10 Gaps: 2

US-09-782-953-7 (1-198) x US-09-782-953-14 (1-2212)

QY 1 MetAaPheArGAspPheSerTyraAnPheSerSerLeuIleAlaCyValAlaAnAaP 20  
 DB 34 GTGGACCTGACAGACCTGCCCC-----AGCGCCACCATGCTGTCACCTGACCG 84  
 QY 21 AaPvaIaPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 40  
 DB 85 CCGGTGTCTGTGAGACCGCTGCGCGCCAAATTTGAGATCCCTCTTAAAGAGATGAC 144  
 QY 41 LysAaPThrThrPheGlnTyRPhelYsSerPheLysArgValArgIleAnPheSerAn 60  
 DB 145 AAGGACATCACCTTTACAGATTTTAAAGCTTCAACAGCTCAGAAATAACTTCAGCAAC 204  
 QY 61 ProLeuSerAlaAlaAaPAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 DB 205 CCTTCTCCGACAGACATGCGACGCTCCAGCTGCATAGACTGATTTCTGGGAAAGGA 264  
 QY 81 MetLysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAn 100  
 DB 265 ATGAACTTATATTTCTCTACAGCTTACACATAGAGAGCTCACACCTGCTCCGCAAT 324  
 QY 101 ProAaPLeuGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120  
 DB 325 CCGACACAGACATTTCTGATCTCCCTCCGCTCCGCACTGGAGATGAAACAGTG 384  
 QY 121 GluAaPAlaThrProValIleAaTyraPLeuLeuTyraAlaIleSerLysLeuGlyPro 140  
 DB 385 GAGAGTCCGACCCACATCAATACTATGATCTCTATATGCCATCTCCCAAGCTGGGGCCA 444  
 QY 141 GlyGluLysTyRGlulLeuHisAlaAlaThrAaPProThrProSerValValAlaHisVal 160  
 DB 445 GGGGAAAGATGAAATTCAGACGACGACGACGACCTCCCGCGTGTGTGTCATGTA 504  
 QY 161 CysGluSerAaPngInLysAaGlnGluGluGluGluMetGlyLysMetLysArgPro 180  
 DB 505 TGTGAGCTGATCAAGG---AGGAGGAAGAAAGAGGAATGGAAGATGACGAGACCT 561  
 QY 181 LysAaPLeuIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198  
 DB 562 AAGCCAAATTTATCCAGACCGAGGCGGAGTACAGCCGATCCACCTCAGC 615

RESULT 6  
 US-09-782-953-1  
 / Sequence 1, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHMEL, BEVERLY  
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 / FILE REFERENCE: UTS:674P21  
 / CURRENT APPLICATION NUMBER: US/09/782,953  
 / PRIOR FILING DATE: 2001-02-13  
 / PRIOR APPLICATION NUMBER: 60/216,601  
 / NUMBER OF SEQ ID NOS: 27  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 1  
 / LENGTH: 599  
 / TYPE: DNA  
 / ORGANISM: Mus musculus  
 US-09-782-953-1

Alignment Scores:  
 Pred. No.: 3.22e-97 Length: 599  
 Score: 868.00 Matches: 170  
 Percent Similarity: 98.84% Conservative: 0  
 Best Local Similarity: 98.84% Mismatches: 2  
 Query Match: 83.22% Indels: 1  
 DB: 10 Gaps: 0

US-09-782-953-7 (1-198) x US-09-782-953-1 (1-599)

QY 2 AaPheArGAspPheSerTyraAnPheSerSerLeuIleAlaCyValAlaAnAaP 21  
 DB 83 GATTTTGGACCTTACCTACATTTTAACTTCTGATGCTGTGTGGAAACGATGAT 142  
 QY 22 ValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 41  
 DB 143 GTCTTACGCAAGTGAAGACAGGCGCAATTTGAAATCCCTTTCAGAACATATGACAA 202  
 QY 42 AaPThrThrPheGlnTyRPhelYsSerPheLysArgValArgIleAnPheSerAnPro 61  
 DB 203 GACACACCTTCCAGATTTTAAAGCTTCAACAGTGTCCGATTAACCTTCAGCAACCC 262  
 QY 62 LeuSerAlaAlaAaPAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81  
 DB 263 TTATCTGACGCCATGCTCAGCTCCGCTCCAGACGAGCTTCCTGGGAAAGAAATG 322  
 QY 82 LysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAnPro 101  
 DB 323 AAGTGTATTTGCTCAGACTTACATAGAGAGTTCACACCTGCTCCGCAAT-CCC 381  
 QY 102 AaPLeuGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 121  
 DB 382 GACAAACAGTCTCATCTCCCTCCGCTCTCTCCCTGCTGGTGAACAAAGTGA 441  
 QY 122 AaPAlaThrProValIleAaTyraPLeuLeuTyraAlaIleSerLysLeuGlyProGly 141  
 DB 442 GATGACACCCCGCATTAATTTACATCTTTATATGCAATCTCCAACTGGGGCCAGGA 501  
 QY 142 GlyLysTyRGlulLeuHisAlaAlaThrAaPProThrProSerValValAlaHisVal 161  
 DB 502 CAGAAATGAACTGACATGACGACGACGACACCACTCCCACTGTGTGTCACAGTGT 561  
 QY 162 GluSerAaPngInLysAaGlnGluGluGluGluMetGlyLysMetLysArgPro 173  
 DB 562 GAGAGTGAACAAAGAAATGAGAGAGAGAGAGAG 597

RESULT 7  
 US-09-880-107-3340  
 / Sequence 3340, Application US/09880107  
 / Patent No. US20020142981A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Horne, Darci T.  
 / APPLICANT: Vockley, Joseph G.  
 / APPLICANT: Scherf, Uwe  
 / APPLICANT: Gene Logic, Inc.  
 / TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 / FILE REFERENCE: 44921-5028-NO  
 / CURRENT APPLICATION NUMBER: US/09/880,107  
 / PRIOR FILING DATE: 2001-06-14  
 / PRIOR APPLICATION NUMBER: US 60/211,379  
 / PRIOR FILING DATE: 2000-06-14  
 / PRIOR APPLICATION NUMBER: US 60/237,054  
 / NUMBER OF SEQ ID NOS: 10-02  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 3340  
 / LENGTH: 2173  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833  
 US-09-880-107-3340



## Alignment Scores:

Pred. No.: 6.87e-94 Length: 2173  
 Score: 847.50 Matches: 162  
 Percent Similarity: 97.63% Conservative: 3  
 Best Local Similarity: 95.86% Mismatches: 3  
 Query Match: 81.26% Indels: 1  
 DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x US-09-880-107-3340 (1-2173)

QY 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysPheThrPheGlnTyrPheLys 49  
 DB 58 GCCAAATTTAGTCCCTTTTGGACGTATGACAGGACATCACCTTTTCAGTATTTTAA 117  
 QY 50 SerPheLysArgValaArgileAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
 DB 118 AGCTTCAACAGGTGAGTAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGCTC 177  
 QY 70 ArgLeuHisLysThrGluPheGluGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89  
 DB 178 CAGTGCATAAGACTGAGTTTCTGGGAAGGAAATGAAGTTATATTTTGTCTCAGACCTTA 237  
 QY 90 HisLysGlySerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109  
 DB 238 CACATAGGAGCTCACACCTGGCTCCGCCAAATCCACAGCAGATTTCTGATCTCCCT 297  
 QY 110 ProLysProProValGlyTTPLeGlnValGluAspAlaThrProValIleAsnTyr 129  
 DB 298 CCCGCCCTCTCCGAGTGGATGAAACAAGTGGAGATCGGACCCGAGTCATAAACTAT 357  
 QY 130 AspLeuLeuTyrAlaLysSerLysLeuGlyProGlyLysTyrGluLeuHisAlaAla 149  
 DB 358 GATCTCTTATATGCTATCTCCAGCTGGGCGCAGGGGAAAGATGAATTTGCACGACGCG 417  
 QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
 DB 418 ACTGACACCACTCCACGCTGGTGTCCATGTATGTGAGATGATCAAGAG---AAGGAG 474  
 QY 170 GluGluGluMetGluArgMetLysArgProLysPheLysLeuGlnThrArgArg 189  
 DB 475 GAAGAAGAGGAAATGGAAGAATGAGGAGACCTTAAGCCAAATTTATCCAGACGAGGAG 534  
 QY 190 ProGluTyrThrProLysHisLeuSer 198  
 DB 535 CCGGAGTACACGCGATCCACCTCAGC 561

## RESULT 8

US-09-925-302-347  
 ; Sequence 347, Application US/09925302  
 ; Patent No. US20020044941A1

## GENERAL INFORMATION:

; APPLICANT: ROSEN ET AL.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA104  
 ; CURRENT APPLICATION NUMBER: US/09/925,302  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 896  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 347  
 ; LENGTH: 2358  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

US-09-925-302-347

## Alignment Scores:

Pred. No.: 1.06e-91 Length: 2358  
 Score: 830.00 Matches: 168  
 Percent Similarity: 88.89% Conservative: 8  
 Best Local Similarity: 84.85% Mismatches: 18

Query Match: 79.58% Indels: 5  
 DB: 10 Gaps: 2  
 US-09-782-953-7 (1-198) x US-09-925-302-347 (1-2358)  
 QY 1 MetAspPheArgAspPheSerTyrAsnPheSerLeuIleAlaCysValAlaAsnAsp 20  
 DB 123 GTGGACCTGCAGGACCTGCC-----AGGCCACCATGCTGTCCCTGACCTGGACCGG 173  
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 DB 174 CGCGTGTCTGTGACCGCGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGAGGTATGAC 233  
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 DB 234 AGGACATCACCTTTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGAAATAAATTCAGCAAC 293  
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 DB 294 CCTTCTCCGACGACAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGAAAGGAA 353  
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100  
 DB 354 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACTGGCTCCGCA-AAT 412  
 QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTTPLeGlnVal 120  
 DB 413 CCAGACAGCAGGTTTCTGATCTCCCTCTCCGCTCTCCGSCAGTGGGATGGAACAAGTG 472  
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaLysSerLysLeuGlyPro 140  
 DB 473 GAAGATGGACCCAGTCATAAACTATGATCTCTTATATGCTATCTCCAGCTGGGCGCA 532  
 QY 141 GlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160  
 DB 533 GGGGAAAGATGAATTCAGCAGCAGCTGACACCACTCCAGCGGTGGTGTCCATGTA 592  
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180  
 DB 593 TGTAGAGTGTCAAGAG---AAGGAGGAAGAGGAATGGAAGAATGAGGAGACCT 649  
 QY 181 LysProLysIleGlnThrArgArgProGluTyrThrProLysHisLeuSer 198  
 DB 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACGCGATCCACCTCAGC 703

## RESULT 9

US-09-782-953-8  
 ; Sequence 8, Application US/09782953  
 ; Patent No. US20020150953A1

## GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; FILE REFERENCE: UTSID:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 594  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(591)

US-09-782-953-8

## Alignment Scores:

Pred. No.: 1.89e-68 Length: 594  
 Score: 632.50 Matches: 121



; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 17  
 ; LENGTH: 3184  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (205)..(780)  
 US-09-782-953-17

## Alignment Scores:

Pred. No.: 2,686-65 Length: 3184  
 Score: 615.50 Matches: 119  
 Percent Similarity: 75.53% Conservative: 23  
 Best Local Similarity: 63.30% Mismatches: 37  
 Query Match: 59.01% Indels: 9  
 DB: 10 Gaps: 2

US-09-782-953-7 (1-198) x US-09-782-953-17 (1-3184)

Qy 11 SerSerLeuIleAlaCysValAlaAsnAspValPheSerGluSerGluThrArgAla 30  
 Db 220 TCCACTCTGGTGGCTGTGGTGGATCTCAGAGTCTTTACCAATCAGAGGTTAAGGAA 279  
 Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
 Db 280 AAATTGGGGGACTGTTTCGACTATGATGACTGTGTGAGCTTCCAGCTATTAAAGT 339  
 Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
 Db 340 TTCAGAGCTGTCGGTATAAACTTCAGCAATCTTAAATCTGCAGCCCGAGCTAGATAGAG 399  
 Qy 71 LeuHisLysThrGluPheLeuGluLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
 Db 400 CTTCAAGAACCAATTACAGGGGAAAAATTAAAGCTCTACTTTGCACAGGTTTCAGACT 459  
 Qy 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106  
 Db 460 CCAGAGACAGATGAGACAACTGCTTGGCTCCACCCAGCTGCCAACAGTTTCTC 519  
 Qy 107 IleSerProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProVal 126  
 Db 520 ATCTGCCCCCTTCTCTCCACCTGTTAGCTGGCAGCCCATCAACGATGCCACGCCAGTC 579  
 Qy 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146  
 Db 580 CTCACCTATGACCTCTCTATGCTGTGGCCAACTAGGACCGAGAGAGTATGAGCTC 639  
 Qy 147 HisAlaIleThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166  
 Db 640 CATGACAGGACTGAGTCCACCCCAAGTGTGCTGCAGCTGGCCAGCTGCATAGAG 699  
 Qy 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysLeuLeuGln 186  
 Db 700 GAAGAGAGGAGCCCAAG-----ACTTCCCCAAAGGCAAAATCATCCAA 744  
 Qy 187 ThrArgArgProGluTyrThrPro 194  
 Db 745 ACTCGGCTCTGGCTGCCACCC 768

## RESULT 12

US-09-782-953-20  
 ; Sequence 20, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674PZ1  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 20  
 ; LENGTH: 828  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (23)..(745)  
 US-09-782-953-20

## Alignment Scores:

Pred. No.: 2,386-63 Length: 828  
 Score: 592.50 Matches: 120  
 Percent Similarity: 71.28% Conservative: 19  
 Best Local Similarity: 61.54% Mismatches: 43  
 Query Match: 56.81% Indels: 13  
 DB: 10 Gaps: 3

US-09-782-953-7 (1-198) x US-09-782-953-20 (1-828)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20  
 Db 140 ATGGATTTAAGTATGCTGCT-----ACCTCATTCTTTGCTTGCAGCTCATGAA 190  
 Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
 Db 191 GCAGTGTGTGAGGCGCAGAGCAGAGGAAAGATTGAAGCACTCTTACCATCTATGAT 250  
 Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
 Db 251 GACCAGGTTTACTTTTTCAGCTGTTTAAAGCTTTAAGAGTTCAGATTAATTTTCAGAAA 310  
 Qy 61 ProLeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
 Db 311 CTTGAAGCGGCGCAGCAGAGCGGAAATAGAACTCCACGAAACAGACTTCAATGGCGAGAAG 370  
 Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGly-----SerSerHisLeu 96  
 Db 371 CTAAAGCTATATTTTGCACAGTGCAGATGTCGCGGGAAGTCCGCGCAAGTCTCTATCTC 430  
 Qy 97 AlaProProAsnProAspLysGlnPheLeuIleSerProProAlaSerProProValGly 116  
 Db 431 CTGCGCGCCCGCAGCTGTCAAGCAGTTCCTCATCTCCCTCCAGCTCTCCCGCAGTGGG 490  
 Qy 117 TrpLysGlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSer 136  
 Db 491 TGGAGCAGACGAGAGAGATGCGATGCTGTTTATAAATTTATGATTTACTCTGTGCTTTCC 550  
 Qy 137 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaIleThrAspProThrProSerVal 156  
 Db 551 AAATTGGGACCGAGAGAAATATGAATCTCCCGGGAACGAGTGCACACCCAGCGTG 610  
 Qy 157 ValValHisValCysGluSerAspGlnGluAsnGluGluGluGluMetGluArg 176  
 Db 611 GTGGTTCATGTCTGTGAAGTGAACCTGAAGAGGAGGAGGAG----- 652  
 Qy 177 MetLysArgProLysProLysIleIleGlnThrArgProGlu 191  
 Db 653 ACAAAAACCCCAACAGAAAAATTGCCACAGACAGAGCGCCCGAC 697

## RESULT 13

US-09-782-953-23  
 ; Sequence 23, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 FILE REFERENCE: US 674P21  
 CURRENT APPLICATION NUMBER: US/09/782,953  
 CURRENT FILING DATE: 2001-02-13  
 PRIOR APPLICATION NUMBER: 60/216,601  
 PRIOR FILING DATE: 2000-07-07  
 NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO: 23  
 LENGTH: 720  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (2)...(637)  
 US-09-782-953-23

Alignment Scores:  
 Pred. No.: 2,578-63 Length: 720  
 Score: 591.50 Matches: 119  
 Percent Similarity: 72.25% Conservative: 19  
 Best Local Similarity: 62.30% Mismatches: 38  
 Query Match: 56.71% Indels: 15  
 DB: Gaps: 3

US-09-782-953-7 (1-198) x US-09-782-953-23 (1-720)

QY 1 MetAaPheArgAaPheSerTyraenPheSerSerLeuIleAlaCyValAlaAaenAp 20  
 DB 62 ATGATTTTAAAGTATTCCTGCTT-----ACCTGACTTTTGGCTTGACGCTCATGAA 112  
 QY 21 AaPvaIaPheSerGluSerGluThraGlaAlaAaPheGluSerLeuPheArgThrTyraP 40  
 DB 113 GCAGGTGTTGAGCAGCAGAGCAGAGAAAGATTGAAAGACTTTCACCATCTATGAT 172  
 QY 41 LysAaPThrThraPheGlnTyraPheLysSerPheLysArgValArgIleAaPheSerAa 60  
 DB 173 GACCAAGTTACTTTCAGCTGCTTTAAAGCTTTAGAAAGCTCAGATTAATTTCAGCAAA 232  
 QY 61 ProLeuSerAlaAlaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaP 80  
 DB 233 CCGAAGCGGCAAGAGCGGCAATGAACTCCACGAAACACACTTCATGAGCGGCAAG 292  
 QY 81 MetLysLeuTyraPheAaGlnThraPheLysIleGlySerSerIleuAaPvaIaAaPvaI 100  
 DB 293 CTAAAGCTATATTGTCAGAG-----TCTATCTCCGCGGCGCCCAAG 334  
 QY 101 ProAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaI 120  
 DB 335 CCGTCAAGAGCTTCCCTCATCTCCCTCCAGGCTCCCTCCAGGGGGTGAAGCAGAGC 394  
 QY 121 GluAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaI 140  
 DB 395 GAAGATCCGAGCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 454  
 QY 141 GlyGlnLysTyraPheLysIleAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaP 160  
 DB 455 GGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 514  
 QY 161 CysGluSerAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaP 180  
 DB 515 TGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 556  
 QY 181 LysPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaPvaIaAaP 191  
 DB 557 AAACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 589  
 RESULT 14  
 US-09-864-761-10388/C  
 ; Sequence 10388, Application US/09864761  
 ; Patent No. US20020048763A1

GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: Aecmica-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 10388  
 LENGTH: 412  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AP000054.1  
 OTHER INFORMATION: EXPRESSED IN HEILA, SIGNAL = 0.6  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67  
 US-09-864-761-10388

Alignment Scores:  
 Pred. No.: 2,756-26 Length: 412  
 Score: 285.50 Matches: 58  
 Percent Similarity: 82.43% Conservative: 3  
 Best Local Similarity: 78.38% Mismatches: 10  
 Query Match: 27.37% Indels: 3  
 DB: Gaps: 1

US-09-782-953-7 (1-198) x US-09-864-761-10388 (1-412)

```

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
Db 410 GCCAAATTTGAGTCCCTCTTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAA 351
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 350 AGCTTCAACGAGTCAAGATAAATTCAGCAACCCCTTCTCCGACGACAGATCCAGGCTC 291
QY 70 ArgLeuHisLysThrGluPheLeuGlyGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 290 CAGCTGCATGAAGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTCTCAGGTGAGT 231
QY 90 HisIle-----GlySerSerHisLeuAlaProAsn 100
Db 230 TGGGTTTCATTGCTATGATGCTTCTCCCTCCCTCCCTCCCTCC 189

RESULT 15
US-09-864-761-2064/c
; Sequence 2064, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2064
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: MAP TO AP000122.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
US-09-864-761-2064

```

```

Alignment Scores:
Pred. No.: 3,1e-26 Length: 446
Score: 285.50 Matches: 58
Percent Similarity: 82.43% Conservative: 3
Best Local Similarity: 78.38% Mismatches: 10
Query Match: 27.37% Indels: 3
DB: 10 Gaps: 1

```

US-09-782-953-7 (1-198) x US-09-864-761-2064 (1-446)

```

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
Db 444 GCCAAATTTGAGTCCCTCTTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAA 385
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 384 AGCTTCAACGAGTCAAGATAAATTCAGCAACCCCTTCTCCGACGACAGATCCAGGCTC 325
QY 70 ArgLeuHisLysThrGluPheLeuGlyGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 324 CAGCTGCATGAAGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTCTCAGGTGAGT 265
QY 90 HisIle-----GlySerSerHisLeuAlaProAsn 100
Db 264 TGGGTTTCATTGCTATGATGCTTCTCCCTCCCTCCCTCCCTCC 223

```

Search completed: December 15, 2002, 00:06:25  
Job time : 38.0556 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:09:05 ; Search time 1042.77 Seconds  
(without alignments)  
14442.598 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gagggtcaaggaaactcca.....gaggagggaaggagagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/1/pna/PCTUS COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US05 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US07 COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US081 COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US082 COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US083 COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pna/US084 COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pna/US085 COMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pna/US086 COMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pna/US087 COMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pna/US088 COMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pna/US089 COMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pna/US090 COMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pna/US091 COMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pna/US092 COMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pna/US093 COMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pna/US094 COMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pna/US095 COMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pna/US095A COMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pna/US095B COMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pna/US095C COMB.seq.\*
- 23: /cgn2\_6/ptodata/1/pna/US095D COMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pna/US096 COMB.seq.\*
- 25: /cgn2\_6/ptodata/1/pna/US096B COMB.seq.\*
- 26: /cgn2\_6/ptodata/1/pna/US096C COMB.seq.\*
- 27: /cgn2\_6/ptodata/1/pna/US096D COMB.seq.\*
- 28: /cgn2\_6/ptodata/1/pna/US096E COMB.seq.\*
- 29: /cgn2\_6/ptodata/1/pna/US097A COMB.seq.\*
- 30: /cgn2\_6/ptodata/1/pna/US097B COMB.seq.\*
- 31: /cgn2\_6/ptodata/1/pna/US097C COMB.seq.\*
- 32: /cgn2\_6/ptodata/1/pna/US098 COMB.seq.\*
- 33: /cgn2\_6/ptodata/1/pna/US098A COMB.seq.\*
- 34: /cgn2\_6/ptodata/1/pna/US098B COMB.seq.\*
- 35: /cgn2\_6/ptodata/1/pna/US099A COMB.seq.\*
- 36: /cgn2\_6/ptodata/1/pna/US099B COMB.seq.\*
- 37: /cgn2\_6/ptodata/1/pna/US099C COMB.seq.\*
- 38: /cgn2\_6/ptodata/1/pna/US099D COMB.seq.\*
- 39: /cgn2\_6/ptodata/1/pna/US100A COMB.seq.\*
- 40: /cgn2\_6/ptodata/1/pna/US100B COMB.seq.\*
- 41: /cgn2\_6/ptodata/1/pna/US101A COMB.seq.\*
- 42: /cgn2\_6/ptodata/1/pna/US101B COMB.seq.\*
- 43: /cgn2\_6/ptodata/1/pna/US102A COMB.seq.\*
- 44: /cgn2\_6/ptodata/1/pna/US102B COMB.seq.\*

- 44: /cgn2\_6/ptodata/1/pna/US6000 COMB.seq.\*
- 45: /cgn2\_6/ptodata/1/pna/US6001 COMB.seq.\*
- 46: /cgn2\_6/ptodata/1/pna/US6002 COMB.seq.\*
- 47: /cgn2\_6/ptodata/1/pna/US6003 COMB.seq.\*
- 48: /cgn2\_6/ptodata/1/pna/US6004 COMB.seq.\*
- 49: /cgn2\_6/ptodata/1/pna/US6005 COMB.seq.\*
- 50: /cgn2\_6/ptodata/1/pna/US6006 COMB.seq.\*
- 51: /cgn2\_6/ptodata/1/pna/US6007 COMB.seq.\*
- 52: /cgn2\_6/ptodata/1/pna/US6008 COMB.seq.\*
- 53: /cgn2\_6/ptodata/1/pna/US6009 COMB.seq.\*
- 54: /cgn2\_6/ptodata/1/pna/US6010 COMB.seq.\*
- 55: /cgn2\_6/ptodata/1/pna/US6011 COMB.seq.\*
- 56: /cgn2\_6/ptodata/1/pna/US6012 COMB.seq.\*
- 57: /cgn2\_6/ptodata/1/pna/US6013 COMB.seq.\*
- 58: /cgn2\_6/ptodata/1/pna/US6014 COMB.seq.\*
- 59: /cgn2\_6/ptodata/1/pna/US6015 COMB.seq.\*
- 60: /cgn2\_6/ptodata/1/pna/US6016 COMB.seq.\*
- 61: /cgn2\_6/ptodata/1/pna/US6017 COMB.seq.\*
- 62: /cgn2\_6/ptodata/1/pna/US6018 COMB.seq.\*
- 63: /cgn2\_6/ptodata/1/pna/US6019 COMB.seq.\*
- 64: /cgn2\_6/ptodata/1/pna/US6020 COMB.seq.\*
- 65: /cgn2\_6/ptodata/1/pna/US6021 COMB.seq.\*
- 66: /cgn2\_6/ptodata/1/pna/US6022 COMB.seq.\*
- 67: /cgn2\_6/ptodata/1/pna/US6023 COMB.seq.\*
- 68: /cgn2\_6/ptodata/1/pna/US6024 COMB.seq.\*
- 69: /cgn2\_6/ptodata/1/pna/US6025 COMB.seq.\*
- 70: /cgn2\_6/ptodata/1/pna/US6026 COMB.seq.\*
- 71: /cgn2\_6/ptodata/1/pna/US6027 COMB.seq.\*
- 72: /cgn2\_6/ptodata/1/pna/US6028 COMB.seq.\*
- 73: /cgn2\_6/ptodata/1/pna/US6029 COMB.seq.\*
- 74: /cgn2\_6/ptodata/1/pna/US6030 COMB.seq.\*
- 75: /cgn2\_6/ptodata/1/pna/US6031 COMB.seq.\*
- 76: /cgn2\_6/ptodata/1/pna/US6032 COMB.seq.\*
- 77: /cgn2\_6/ptodata/1/pna/US6033 COMB.seq.\*
- 78: /cgn2\_6/ptodata/1/pna/US6034 COMB.seq.\*
- 79: /cgn2\_6/ptodata/1/pna/US6035 COMB.seq.\*
- 80: /cgn2\_6/ptodata/1/pna/US6036 COMB.seq.\*
- 81: /cgn2\_6/ptodata/1/pna/US6037 COMB.seq.\*
- 82: /cgn2\_6/ptodata/1/pna/US6038 COMB.seq.\*
- 83: /cgn2\_6/ptodata/1/pna/US6039 COMB.seq.\*
- 84: /cgn2\_6/ptodata/1/pna/US6040 COMB.seq.\*
- 85: /cgn2\_6/ptodata/1/pna/US6041 COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599	100.0	599	30	US-09-782-953-1
2	554	92.5	592	28	US-09-704-424-23327
3	504.4	84.2	597	30	US-09-782-953-5
4	500	83.5	538	28	US-09-704-424-23248
5	427.8	71.4	597	18	US-60-360-207-8520
6	427.4	71.4	613	18	US-09-401-645-4154
7	426.2	71.2	597	22	US-09-575-580B-2
8	426.2	71.2	597	30	US-09-782-953-2
9	413.8	69.1	1880	18	US-09-432-241A-4522
10	413.8	69.1	2331	30	US-09-782-953-11
11	413.8	69.1	2340	42	US-10-247-671-100
12	413.8	69.1	2340	76	US-60-323-784-100
13	413.8	69.1	2346	1	PCT-US02-02176-1165
14	413.8	69.1	2346	1	PCT-US02-23766-4321
15	413.8	69.1	2355	38	US-10-044-090-255
16	413.8	69.1	2355	39	US-10-084-817-36
17	413.8	69.1	2355	40	US-10-116-802-289
18	413.8	69.1	2355	70	US-60-260-483-255
19	413.8	69.1	2355	72	US-60-281-593-289
20	412.2	68.8	580	25	US-09-544-863-7097
21	410	68.4	456	18	US-09-436-762A-33053

22 405.8 67.7 594 25 US-09-652-355-3836 Sequence 1836, Ap  
23 405 67.6 440 18 US-09-436-762A-28336 Sequence 28336, A  
24 404.8 67.6 2348 63 US-09-195-106-60 Sequence 60, Ap1  
25 402.8 67.2 581 28 US-09-710-280-1577 Sequence 1577, Ap1  
26 400.8 66.9 567 20 US-09-539-802A-533 Sequence 533, App  
27 399.6 66.7 593 18 US-09-401-645-9836 Sequence 3836, Ap  
28 394.4 65.8 593 29 US-09-726-787-1970 Sequence 1970, Ap  
29 384 64.1 578 28 US-09-710-280-1000 Sequence 1000, Ap  
30 377 62.9 401 18 US-09-436-762A-16620 Sequence 16620, A  
31 374.4 62.5 972 17 US-09-359-067-31161 Sequence 31161, A  
32 373 62.3 397 18 US-09-436-762A-18907 Sequence 18907, A  
33 370.6 61.9 518 25 US-09-652-128-4690 Sequence 4690, Ap  
34 369.4 61.7 586 29 US-09-726-787-1963 Sequence 1963, Ap  
35 369 61.6 393 18 US-09-436-762A-31396 Sequence 31396, A  
36 368 61.4 581 25 US-09-539-802A-842 Sequence 842, App  
37 360.8 60.2 543 20 US-09-644-869-7149 Sequence 7149, Ap  
38 360 60.1 566 25 US-09-644-869-6053 Sequence 6053, Ap  
39 359.8 60.1 540 29 US-09-726-787-1098 Sequence 1098, Ap  
40 357.8 59.7 2358 1 PCT-US00-05918-347 Sequence 347, App  
41 357.8 59.7 2358 34 US-09-925-302-347 Sequence 347, App  
42 357.4 59.7 524 25 US-09-652-123-3252 Sequence 3252, App  
43 357.4 59.7 562 20 US-09-539-802A-2692 Sequence 2692, Ap  
44 356.4 58.5 552 28 US-09-710-280-1359 Sequence 1359, Ap  
45 350.6 58.5 563 28 US-09-716-475-4375 Sequence 4375, Ap

## ALIGNMENTS

RESULT 1  
US-09-782-953-1  
Sequence 1, Application US/09782953

GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: US/09-674P21  
CURRENT APPLICATION NUMBER: US/09/782,953  
CURRENT FILING DATE: 2001-02-13  
PRIORITY APPLICATION NUMBER: 60/216,601  
PRIORITY FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 599  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-782-953-1

Query Match 100.0%; Score 599; DB 30; Length 599;  
Best Local Similarity 100.0%; Pred. No. 5,5e-172;  
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGTGCAAGAACTCCAGCTTGGCTTGAAGAGCCAGTGTCTTAAAGCT 60  
DB 1 GAGGTGCAAGAACTCCAGCTTGGCTTGAAGAGCCAGTGTCTTAAAGCT 60  
QY 61 CTGCCCCGTAAGAAAGAGAAATATTTAGGACTTAACTAATTTAGTCCCTGAT 120  
DB 61 CTGCCCCGTAAGAAAGAGAAATATTTAGGACTTAACTAATTTAGTCCCTGAT 120  
QY 121 TGGTGTGTGCAAAAGATGATCTTCAAGCAAGTGAAGCCAGGCAAAATTGATC 180  
DB 121 TGGTGTGTGCAAAAGATGATCTTCAAGCAAGTGAAGCCAGGCAAAATTGATC 180  
QY 181 CTTCTTGAAGCATATGACAAAGACACCACTTCCAGTATTTAAGAGCTTCAAGCTGT 240  
DB 181 CTTCTTGAAGCATATGACAAAGACACCACTTCCAGTATTTAAGAGCTTCAAGCTGT 240  
QY 241 CCGGATAACTTGCAAGACCCCTTATCTGAGCCGATGCCAGGCTGGGCTGCAAGAC 300  
DB 241 CCGGATAACTTGCAAGACCCCTTATCTGAGCCGATGCCAGGCTGGGCTGCAAGAC 300

QY 301 CGAGTCTCGGGAGAGAAATGATGTAATTTTGTCTCAAGCTTTACATAGAACTTC 360  
DB 301 CGAGTCTCGGGAGAGAAATGATGTAATTTTGTCTCAAGCTTTACATAGAACTTC 360  
QY 361 ACACCTGCTCCCGCATTCGCCAACAACAGTTCTCTATCCCTCGGGCTCTCCCTCC 420  
DB 361 ACACCTGCTCCCGCATTCGCCAACAACAGTTCTCTATCCCTCGGGCTCTCCCTCC 420  
QY 421 GTTGGCTGGAACAAGTAGAAGATGCCACCCCTCTAATAATTCATGATCTTTATATGCC 480  
DB 421 GTTGGCTGGAACAAGTAGAAGATGCCACCCCTCTAATAATTCATGATCTTTATATGCC 480  
QY 481 ATTCCTCAAGCTGGGCGCCAGAGAGAGATGATGACTGATGAGCGACACCACTTCC 540  
DB 481 ATTCCTCAAGCTGGGCGCCAGAGAGAGATGATGACTGATGAGCGACACCACTTCC 540  
QY 541 AGTGTGTGTCTCCACGTGTGTGAGAGTGAACCAAGAAATGAGGAGAAAGAGAT 599  
DB 541 AGTGTGTGTCTCCACGTGTGTGAGAGTGAACCAAGAAATGAGGAGAAAGAGAT 599

RESULT 2  
US-09-704-424-23327  
Sequence 23327, Application US/09704424

GENERAL INFORMATION:  
APPLICANT: Watson, James  
TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed  
FILE REFERENCE: 11000,1045U  
CURRENT APPLICATION NUMBER: US/09/704,424  
CURRENT FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 36881  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23327  
LENGTH: 592  
TYPE: DNA  
ORGANISM: Mouse  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(592)  
OTHER INFORMATION: n = A,T,C or G  
US-09-704-424-23327

Query Match 92.5%; Score 554; DB 28; Length 592;  
Best Local Similarity 99.3%; Pred. No. 3.3e-158;  
Matches 587; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 11 GGAACCTCCAGCTTGGGCTTGAAGAGAGAGAGTGTGTTAAGGCTTGCCCGCTG 70  
DB 1 GGAACCTCCAGCTTGGGCTTGAAGAGAGAGAGTGTGTTAAGGCTTGCCCGCTG 70  
QY 71 AAAAAGCAAGATG-ATTTAGGCACTTATGACAAATTTAGCTCCCTGATGCTTGTGT 129  
DB 60 AAAAAGCAAGATGATTTTAGGCACTTATGACAAATTTAGCTCCCTGATGCTTGTGT 119  
QY 130 GGCACAAGATGATGCTTCAAGCAAGTGAAGACAGGCGCAATTTGATCCCTTTGAG 189  
DB 120 GGCACAAGATGATGCTTCAAGCAAGTGAAGACAGGCGCAATTTGATCCCTTTGAG 179  
QY 190 AACATATGCAAGGACACCACTTCCAGTATTTAAGAGCTTCAAAAGTGTCCGATAA 249  
DB 180 AACATATGCAAGGACACCACTTCCAGTATTTAAGAGCTTCAAAAGTGTCCGATAA 239  
QY 250 CTTTCAAGCAACCTTATATGCAAGCCAGTCCAGGCTGGGCTGCAAGACCGAGTCTCT 309  
DB 240 CTTTCAAGCAACCTTATATGCAAGCCAGTCCAGGCTGGGCTGCAAGACCGAGTCTCT 299  
QY 310 GGGGAAGAAATGAAGTTTATTTTGTCTCAAGCTTATACATAGAAATTTCACTGGC 369  
DB 300 GGGGAAGAAATGAAGTTTATTTTGTCTCAAGCTTATACATAGAAATTTCACTGGC 359  
QY 370 TCCG-CCATCCGCAACAAAGTTCTCTATCCCTCGGGCTCTCTCCGCTGGCTG 428





```

; Sequence 8520, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 8520
; LENGTH: 597
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-8520

```

Query Match 71.4%; Score 427.8; DB 80; Length 597;

Best Local Similarity 93.3%; Pred. No. 1,4e-119; Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

```

QY 110 AGCTCCGATGCTGTGTGGCAACGATGATGCTTCAGCAAAAGTGAACCAAGGCC 169
DB 31 AGGCCACCATGCTGCTGCCACCTGACCCGCGCTGTTGCTGAGCGCTGTGCCGGCC 90
QY 170 AATTGAAATCCCTCTTCAGAACATATGACAGACACCACTTCCAGTATTTTAAGAC 229
DB 91 AATTGAAATCCCTCTTCAGAACATATGACAGACACCACTTCCAGTATTTTAAGAC 150
QY 230 TTCAAAAGTGTGGGATTAATCTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTGGG 289
DB 151 TTCAAAAGTGTGGGATTAATCTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTGGG 210
QY 290 CTCACACAGCCAGTCTCTGGGAAAGAAATGAAATGATTTGATTTTCTGAGCTTTAC 349
DB 211 CTCACACAGCCAGTCTCTGGGAAAGAAATGAAATGATTTGATTTTCTGAGCTTTAC 270
QY 350 ATAGGAAGTTCAACCTGCTGCTCG-CCATCCCGACAAAGTTCTCATCTCCCTCCG 408
DB 271 ATAGGAAGTTCAACCTGCTGCTCGCCCAATCCGACAAAGTTCTCATCTCCCTCCG 330
QY 409 GCCTCTCCCTCCGCTGCTGCTGAAACAGTAGAAGATCCACCCCGCATTAATTAAGAT 468
DB 331 GCCTCTCCCTCCGCTGCTGCTGAAACAGTAGAAGATCCACCCCGCATTAATTAAGAT 390
QY 469 CTTTATATGCACTTCTCAAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGAT 528
DB 391 CTTTATATGCACTTCTCAAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGAT 450
QY 529 GACACCACTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 588
DB 451 GACACCACTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
QY 589 GAGGAGAGAT 599
DB 511 GAGGAGAGAT 521

```

# RESULT 6

US-09-401-645-4194

```

; Sequence 4194, Application US/09401645
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Diselano, Peter
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: RAT SKELETAL MUSCLE LIBRARY
; FILE REFERENCE: MN98-4194
; CURRENT APPLICATION NUMBER: US/09/401,645
; CURRENT FILING DATE: 1999-09-22
; EARLIER APPLICATION NUMBER: 60/101,346
; EARLIER FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 4253
; SOFTWARE: ParseSeq for Windows Version 3.0
; SEQ ID NO 4194
; LENGTH: 613

```

```

; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(613)
; OTHER INFORMATION: n = A,T,C or G
US-09-401-645-4194

```

Query Match 71.4%; Score 427.4; DB 18; Length 613;  
Best Local Similarity 85.1%; Pred. No. 1,9e-119;  
Matches 504; Conservative 0; Mismatches 84; Indels 4; Gaps 3;

```

QY 9 AAGAACCTTCAGCTGGGCTTGACT--GAGAGAGAGATGCTTCTTAAGCGCTGCCC 66
DB 1 AAGAACCTTCAGCTGGGCTTGACTGAGTGAAGTGTCTTAAGCGCTGCCC 60
QY 67 CGTGAAGAGAGAAAT-ATTTAGGACCTTTAGCTCAATTTAGCTCCCTGATTC 125
DB 61 GTTGAAGAGAGAAATGCTTTTGAAGACTTTTAACTTAACTTAACTTAACTTAACTT 120
QY 126 GTGTGGCAACGATGATGCTTCAGCGAAAGTGAAGACAGGCGCAATTTGAATCCCTCT 185
DB 121 GTGTGGCAACGATGATGCTTCAGCGAAAGTGAAGACAGGCGCAATTTGAATCCCTCT 180
QY 186 TCAGAACATATGACAGACACCACTTCCAGTATTTTAAGAGCTTCAACGTGCGGA 245
DB 181 TCAGAACATATGACAGACACCACTTCCAGTATTTTAAGAGCTTCAACGTGCGGA 239
QY 246 TAACTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGCGGCTGCAGAACCGAGT 305
DB 240 TAACTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGCGGCTGCAGAACCGAGT 299
QY 306 TCTGTGGGAGAAATGAATGATTTTGTCTGACACTTTACATATGAAAGTTCAAC 365
DB 300 TCTGTGGGAGAAATGAATGATTTTGTCTGACACTTTACATATGAAAGTTCAAC 359
QY 366 TGGCTCCCGCAATCCCGCAACAGTTGCTCATCTCCCTCCGCTCTCTCCGCTGG 425
DB 360 TGGCTCCCGCAATCCCGCAACAGTTGCTCATCTCCCTCCGCTCTCTCCGCTGG 419
QY 426 CTGAAACAGTAGAAGATGCTGACCCCGCATTAATTAAGATCTTTATATGATCTC 485
DB 420 CTGAAACAGTAGAAGATGCTGACCCCGCATTAATTAAGATCTTTATATGATCTT 479
QY 486 CAAGCTGGGCGCAGAGAGAGATGAACTGATGACAGCGACAGACCACTCCAGGT 545
DB 480 CAAGCTGGGCGCAGAGAGAGATGAACTGATGACAGCGACAGACCACTCCAGGT 539
QY 546 GGTGTCCACGTTGTGTGAGATGACCAAGAGATGAGAGAGAGAGAGAG 597
DB 540 GGTGTCCACGTTGTGTGAGATGACCAAGAGATGAGAGAGAGAGAGAG 591

```

# RESULT 7

US-09-575-580B-2

```

; Sequence 2, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKee, P.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048-01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-575-580B-2

```

Query Match 71.2%; Score 426.2; DB 22; Length 597;  
Best Local Similarity 93.1%; Pred. No. 4.3e-119;  
Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTGCTGTGCGCAACGATGATGCTTCAGCGGAAGTGAGACGAGGCC 169  
DB 31 AGCGCCACCATCGCTGCGCACCTGACCGCGCGGTGTCGTGGAGCGGCTGTGCGCGGCC 90

QY 170 AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACACCTTCCAGTATTTTAAAGAGC 229  
DB 91 AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACACCTTCCAGTATTTTAAAGAGC 150

QY 230 TTCAAACGCTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCCG 289  
DB 151 TTCAAACGCTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCCG 210

QY 290 CTGCAACAAGCCGAGTTCCTGGGGAAGGAAATGAAGTTGTTATTTTGTCTCAGACTTTACAC 349  
DB 211 CTGCAACAAGCCGAGTTCCTGGGGAAGGAAATGAAGTTGTTATTTTGTCTCAGACTTTACAC 270

QY 350 ATAGGAAGTTTACACCTGGCTCCG-CCAAATCCCGAACAACAGTTTCTCATCTCCCTCCG 408  
DB 271 ATAGGAAGTTTACACCTGGCTCCG-CCAAATCCCGAACAACAGTTTCTCATCTCCCTCCG 330

QY 409 GCCTCTCTCCGCTGGTGGAAACAAGTAGAAGTGCACCCCGCTCATTAATTTACGAT 468  
DB 331 GCCTCTCTCCGCTGGTGGAAACAAGTAGAAGTGCACCCCGCTCATTAATTTACGAT 390

QY 469 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAAATGAACTGCATGCAGCGACA 528  
DB 391 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAAATGAACTGCATGCAGCGACA 450

QY 529 GACACCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 588  
DB 451 GACCCCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 510

QY 589 GAGGAAGAGAT 599  
DB 511 GAGGAAGAGAT 521

RESULT 8  
US-09-782-953-2  
; Sequence 2, Application US/09782953  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE OF INVENTION: CALCIUM-INDUCED INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: US/09782,953  
; CURRENT APPLICATION NUMBER: 2001-02-13  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(594)  
US-09-782-953-2

Query Match 71.2%; Score 426.2; DB 30; Length 597;  
Best Local Similarity 93.1%; Pred. No. 4.3e-119;  
Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTGCTGTGCGCAACGATGATGCTTCAGCGGAAGTGAGACGAGGCC 169  
DB 31 AGCGCCACCATCGCTGCGCACCTGACCGCGCGGTGTCGTGGAGCGGCTGTGCGCGGCC 90

QY 170 AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACACCTTCCAGTATTTTAAAGAGC 229  
DB 91 AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACACCTTCCAGTATTTTAAAGAGC 150

QY 230 TTCAAACGCTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCCG 289  
DB 151 TTCAAACGCTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCCG 210

QY 290 CTGCAACAAGCCGAGTTCCTGGGGAAGGAAATGAAGTTGTTATTTTGTCTCAGACTTTACAC 349  
DB 211 CTGCAACAAGCCGAGTTCCTGGGGAAGGAAATGAAGTTGTTATTTTGTCTCAGACTTTACAC 270

QY 350 ATAGGAAGTTTACACCTGGCTCCG-CCAAATCCCGAACAACAGTTTCTCATCTCCCTCCG 408  
DB 271 ATAGGAAGTTTACACCTGGCTCCG-CCAAATCCCGAACAACAGTTTCTCATCTCCCTCCG 330

QY 409 GCCTCTCTCCGCTGGTGGAAACAAGTAGAAGTGCACCCCGCTCATTAATTTACGAT 468  
DB 331 GCCTCTCTCCGCTGGTGGAAACAAGTAGAAGTGCACCCCGCTCATTAATTTACGAT 390

QY 469 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAAATGAACTGCATGCAGCGACA 528  
DB 391 CTTTATATGCCATCTCCAGCTGGGGCCAGGAGAGAAATGAACTGCATGCAGCGACA 450

QY 529 GACACCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 588  
DB 451 GACCCCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGAA 510

QY 589 GAGGAAGAGAT 599  
DB 511 GAGGAAGAGAT 521

RESULT 9  
US-09-432-241A-4522  
; Sequence 4522, Application US/09432241A  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A  
; FILE OF INVENTION: HUMAN EPIDERMAL KERATINOCYTE LIBRARY  
; FILE REFERENCE: 1600.1004001  
; CURRENT APPLICATION NUMBER: US/09432,241A  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/106,445  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/107,227  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 60/127,182  
; PRIOR FILING DATE: 1999-03-30  
; PRIOR APPLICATION NUMBER: 60/132,099  
; NUMBER OF SEQ ID NOS: 5041  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4522  
; LENGTH: 1880  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-432-241A-4522

Query Match 69.1%; Score 413.8; DB 18; Length 1880;  
Best Local Similarity 85.1%; Pred. No. 4.4e-115;  
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGCGAGTCGTTCTGTTAAGCGCTGCCCCCGTAAAAAGCAGATGATTTTAGGCACTTTA 98  
DB 89 GACTGCGTGGTCTGTGACGCTTTCACTTAAGAAAGCAAGATGATTTTAGAACTTTA 148

QY 99 GCTCAAAATTTTAGCTCCCTGATTTGTTGTGGCAACAGATGATGTTCTTCAGCGAAAGTG 158  
DB 149 ACTACAGTTTATAGTCTCCCTGATTTGCTGTGGCAACAGTATGATCTTCAGCGAAAGTG 208

QY 159 AGACGAGGCGCAATTTTGAATCCCTCTTCAGAACATATGACAAGGACACCACTTCCAGT 218

Db	209	AAACCGAGGGCCAATTGAGTCCCTCTTTAGACGATGACAGGACATACACTTTGAGT	268
OY	219	ATTTTAAGACCTTCAACGCTGTCGGATTAACCTTAGCAACCCCTTATCTGCAGCCGATG	278
Db	269	ATTTTAAGAGCTTCAACGAGTGAGAAATAACTTAGCAACCCCTTCTCCGACGAATG	328
OY	279	CCAGGCTGGGGCTGCACAGACCGAGTTCTCTGGGCAAGGAAATGAAGTGTAAATTTGTC	338
Db	329	CCAGGCTCCAGCTGCATAAACACTGAGTTTCTGGAAAGGAAATGAAGTGTAAATTTGTC	388
OY	339	AGACTTAAACATAGAGAGTTGCACCTGAGTCCGCGC-AATCCGACAAAGTGTCTCA	397
Db	389	AGACCTTAAACATAGAGAGCTCAACCTGAGTCCGCGCAATTCACACAGAGTTTCTGA	448
OY	398	TCTCCCTCCGGGCTCTCTCCCGTTGGCTGGAACAACTAGAAAGTGCACCCCGTCA	457
Db	449	TCTCCCTCCCGGCTCTCTCCGCAATGGGAATGGAACAACTAGAAAGTGCACCCCGTCA	508
OY	458	TAAATTACGATCTTTTATATGCAATCTCCAACTGGGGCCAGAGAAAGATGTGAATGCG	517
Db	509	TAAACTATGATCTTTTATATGCAATCTCCAACTGGGGCCAGAGAAAGATGTGAATGCG	568
OY	518	ATGCAGCGACGAGACCAACCTCCGAGTGTGTGCTCCAGCGTGTGAGAGTGCACCAAGAGA	577
Db	569	ACGAGCGACGAGACCAACCTCCGAGCTGTGTGCTCAATGATGTGAGAGTGTCAAGAGA	628
OY	578	ATGAGAGAGAGAGGAA 594	
Db	629	AGGAGAGAGAGAGGAA 645	

```

RESULT 10
US-09-782-953-11
: Sequence 11, Application US/09782953
: GENERAL INFORMATION:
: APPLICANT: WILLIAMS, R. SANDERS
: APPLICANT: ROTHERMEL, BEVERLY
: TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
: TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
: FILE REFERENCE: UTSD:674P21
: CURRENT APPLICATION NUMBER: US/09/782,953
: CURRENT FILING DATE: 2001-02-13
: PRIOR APPLICATION NUMBER: 60/216,601
: PRIOR FILING DATE: 2000-07-07
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 11
: LENGTH: 2331
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (144)..(734)
US-09-782-953-11

```

Query Match	Similarity	69.1%	Score 413.8	DB 30	Length 2331
Best Local	Similarity	85.1%	Pred. No. 4.8e-115		
Matches	474	Conservative	0	Mismatches	82
				Indels	1
				Gaps	1

Db	283	ATTTTAAGACCTTCMAACGAGTCAGAAATAAATTTCAGCAACCCCTTCTCCGACGACGATG	342
Oy	279	CCAGCGTCGGGCTGCAACAACCCAGTCTCTGGGGGAAGAAATGAAGTTGATTTTGTCTC	338
Db	343	CCAGGCTCCAGCTGCAATAAGATGAGTTTCTGGGAAAGAAATGAAGTTATATTTTGTCTC	402
Oy	339	AGACTTTACACATGAGAAATTCAACACTGCTCCGCC-ATTCGGACAAACATTTCTCTCA	397
Db	403	AGACTTTACACATGAGAAAGCTCAACCTGTCTCCGCCAAATCCAGACAAAGCATTTCTGA	462
Oy	398	TCTCCCTCCGGGCTCTCCCTCCGGTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCA	457
Db	463	TCTCCCTCCGGGCTCTCCGCGACGTGGATGGAACAAAGTAGAAGATGCCACCCGACTCA	522
Oy	458	TAAATTAAGATCTTTATATGCCATCTCCAAAGCTGGGGCCAGAGAGAAATATGAATGCC	517
Db	523	TAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAATATGAATTTGC	582
Oy	518	ATGCAAGCACAGAACACCACTCTCCAGTGTGTGTGCCAGTGTGTGAGAGAGTACCAAGAGA	577
Db	583	AACGACGACATCGAACCACTCTCCACGCGTGTGTGTCATGATGTGAGAGTGTATCAAGAGA	642
Oy	578	ATGAGGAGAGAGAGGAA 594	
Db	643	AACGAGAGAGAGAGGAA 659	

```

RESULT 11
US-10-247-671-100
Sequence 100: Application US/10247671
GENERAL INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiftman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaserer, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PERL Program
SEQ ID NO 100
LENGTH: 2340
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 5511889CB1
US-10-247-671-100

```

	Query Match	Similarity	Score	413.8	DB	42	Length	2340
	Best Local	Similarity	85.1%	Pred	No	4.8e-115		
	Matches	474	Conservative	0	Mismatches	82	Indels	1
								Gaps
Qy	39	GAGCGAGTCGTTGGTTAAGCGTCGTGCCCCGGTGAAGAAAGCAGAAATGATTTTAGGCACTTTA	98					
Db	110	GACTGCGTGGGCTGTGAGCGCTTTCACCTTAAGAAAGAACAGATGACATTTTAGAACTTTA	169					
Qy	99	GCTAACATTTTACTCCGCGATTGCTGTGTGGCAAGCATGATGCTTCAGCGAAAGTG	158					
Db	170	ACTACAGTTTACTCCGCGATTGCTGTGTGGCAAGCATGATGCTTCAGCGAAAGTG	229					
Qy	159	AGACGAGGCGCAATTGATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGT	218					
Db	230	AAACGAGGCGCAATTGATCCCTCTTCAGACGATATGACAAAGACATCACTTCCAGT	289					
Qy	219	ATTTTAAAGACTTCAAACGTCGCGGATTAACCTTACGCAACCCCTTATCTGACGCGATG	278					
Db	290	ATTTTAAAGACTTCAAACGATGAGATTAATTAACCTTACGCAACCCCTTCTCGACGAGATG	349					





```

Db 530 TAACTATGATCTTATATGCCATCTCCAAGCTGGGCGGGAAGTATGAATTGC 589
QY 518 ATGCAGCGACAGACACCACTCCCAAGTGTGGTCCACGTGTGTGAGAGTGACCAAGAGA 577
Db 590 ACGCAGCGACTGACACCACTCCCAAGTGTGGTCCATGTATGTGAGAGTGATCAAGAGA 649
QY 578 ATGAGGAGGAGAGGAA 594
Db 650 ACGAGGAAGAGAGGAA 666

```

Search completed: December 14, 2002, 19:55:40  
 Job time : 1047.77 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 72.0383 Seconds  
(without alignments)  
9923.385 Million cell updates/sec

Title: US-09-782-953-11

Perfect score: 2331

Sequence: 1 tttttttttccccaggaggt.....999cacagtatgctcttac 2331

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2059.2	88.3	2174	2	US-08-665-040-1
C 2	246.8	10.6	251	4	US-09-602-877A-83
C 3	52.8	2.3	7218	1	US-08-232-463-14
C 4	41	1.8	1770	4	US-08-939-309-7
C 5	40	1.7	40	2	US-08-665-040-9
C 6	39.6	1.7	2251	4	US-08-991-677-11
C 7	38.2	1.6	7218	1	US-08-232-463-14
C 8	38	1.6	327	4	US-09-889-595-7
C 9	38	1.6	5058	4	US-09-889-595-1
C 10	37.6	1.6	5194	1	US-09-599-652-1
C 11	37.6	1.6	5194	2	US-08-642-848-1
C 12	37.6	1.6	5194	4	US-09-264-604-1
C 13	37.2	1.6	6243	2	US-09-056-075-1
C 14	36.8	1.6	289	4	US-09-007-005-17
C 15	36.8	1.6	289	4	US-09-007-005-17
C 16	36.8	1.6	289	4	US-09-244-796-17
C 17	36.8	1.6	289	4	US-09-244-796-17
C 18	36.6	1.6	960	3	US-08-556-419-9
C 19	36.6	1.6	246240	2	US-08-724-394A-20
C 20	36.6	1.6	246240	2	US-08-724-394A-21
C 21	36.6	1.6	246240	2	US-08-724-394A-22
C 22	36.4	1.6	17656	4	US-09-433-579-3
C 23	36.2	1.6	583	4	US-09-404-671-5
C 24	36.2	1.6	583	4	US-09-404-671-7
C 25	36.2	1.6	162450	4	US-09-345-882-1
C 26	35.8	1.5	65042	4	US-09-784-316-3
C 27	35.4	1.5	4285	4	US-09-410-464-1

C 28	35.4	1.5	7425	4	US-09-453-702B-212
C 29	35.2	1.5	1032	4	US-08-482-073-8
C 30	35	1.5	87350	3	US-08-781-891-79
C 31	35	1.5	87543	4	US-09-791-211-3
C 32	35	1.5	246240	2	US-08-724-394A-20
C 33	35	1.5	246240	2	US-08-724-394A-21
C 34	35	1.5	246240	2	US-08-724-394A-22
C 35	34.2	1.5	688	1	US-08-139-937-9
C 36	34.2	1.5	688	5	PCT-US93-11310-9
C 37	34.2	1.5	5063	1	US-08-185-432-1
C 38	34.2	1.5	5093	1	US-08-468-036-23
C 39	34.2	1.5	5093	2	US-08-376-843-23
C 40	34	1.5	8703	4	US-08-961-527-177
C 41	34	1.5	19124	2	US-08-487-826B-13
C 42	33.8	1.5	667	4	US-09-123-912-39
C 43	33.8	1.5	667	4	US-09-643-597-99
C 44	33.8	1.5	2829	3	US-08-851-843A-53
C 45	33.8	1.5	2829	4	US-08-974-549A-220

#### ALIGNMENTS

#### RESULT 1

US-08-665-040-1  
; Sequence 1, Application US/08665040  
; Patent No. 5869318  
; GENERAL INFORMATION:  
; APPLICANT: ESTIVILL PALLEJA, XAVIER  
; APPLICANT: FUENTES, JUAN JOSE  
; APPLICANT: PRITCHARD, MELANIE  
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPlicing PCR" TECHNIQUE,  
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
; TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD  
; TITLE OF INVENTION: FOR CHARACTERIZING IT.  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY  
; STREET: 26 WEST 61ST STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10023

#### COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1 FOR DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,040  
FILING DATE: JUNE 7, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: ES P9501140  
FILING DATE: JUNE 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: JANET I. CORD  
REGISTRATION NUMBER: 33,778  
REFERENCE/DOCKET NUMBER: U010815-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800  
TELEFAX: (212) 246-8959

#### INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: cDNA for mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE: human  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: foetal  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE: gene library of cDNA  
LIBRARY: gene library of cDNA from foetal  
CLONE: BC-17.8-1 and BC-17.8-2  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9  
MAP POSITION: 21q22.1-q22.2  
FEATURE:  
NAME/KEY: cDNA for mRNA, BC-17.8  
LOCATION: 1..2174  
FEATURE:  
NAME/KEY: untranslated 5'  
LOCATION: 1..48  
FEATURE:  
NAME/KEY: coding sequence  
LOCATION: 49..560  
OTHER INFORMATION: Down Syndrome critical  
FEATURE:  
NAME/KEY: DSCR1  
LOCATION: 1..171 PEPTIDES  
IDENTIFICATION METHOD: translation of the  
OTHER INFORMATION: - deduced protein  
OTHER INFORMATION: - proline-rich protein domains  
OTHER INFORMATION: - glutamic acid-rich protein domains  
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains  
FEATURE:  
NAME/KEY: untranslated 3'  
LOCATION: 564..2174  
FEATURE:  
NAME/KEY: 2 poly (A)  
LOCATION: 1541..1546 AND 2132..2137  
US-08-665-040-1  
Query Match 88.3%; Score 2059.2; DB 2; Length 2174;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2001; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
QY 231 GCCAAATTTGAGTCCCTCTTTAGAGCTATGACAGACATCACTTTCAGTATTTTAAAG 290  
DB 58 GCCAAATTTGAGTCCCTCTTTAGAGCTATGACAGACATCACTTTCAGTATTTTAAAG 117  
QY 291 AGCTTCAAAGAGTCAGATTAATTAATTTAGAGACCCCTTCTCCGACAGATGCGAGCTC 350  
DB 118 AGCTTCAAAGAGTCAGATTAATTAATTTAGAGACCCCTTCTCCGACAGATGCGAGCTC 177  
QY 351 CAGCTGATAGACTGAGTTCTGGGAAAGAAATGAAGTTATATTTTGTCAAGACCTTA 410  
DB 178 CAGCTGATAGACTGAGTTCTGGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTA 237  
QY 411 CACATAGGAAGCTCACACCTGGCTCCGCCAATTCAGACAAAGAGTTTCTGATCTCCCT 470  
DB 238 CACATAGGAAGCTCACACCTGGCTCCGCCAATTCAGACAAAGAGTTTCTGATCTCCCT 297  
QY 471 CCGGCTCTCCGCGAGTGGAGTGAACAAGTGAAGATGAGACCCAGTCATTAATAT 530  
DB 298 CCGGCTCTCCGCGAGTGGAGTGAACAAGTGAAGATGAGACCCAGTCATTAATAT 357  
QY 531 GATCTCTATATGTCATCTCCAGCTGGGGCCAGGGGAAAAGTATGTAATGACGACGCG 590  
DB 358 GATCTCTATATGTCATCTCCAGCTGGGGCCAGGGGAAAAGTATGTAATGACGACGCG 417  
QY 591 ACTGACACCACTCCAGGCTGTGTGTCATATGTCAGAGTGTCAAGGAAGAGAGAA 650  
DB 418 ACTGACACCACTCCAGGCTGTGTGTCATATGTCAGAGTGTCAAGGAAGAGAGAA 477  
QY 651 GAAAGAGAAATGGAAGAAATGAGAGACCTAAGCCAAATTTATCCAGACGAGAGCG 710  
DB 478 GAAAGAGAAATGGAAGAAATGAGAGACCTAAGCCAAATTTATCCAGACGAGAGCG 537

QY 711 GAGTACACGCCGATCCACTCAGCTGAACCTGGACCGGAGAGAGACGATCCCAATCAT 770  
DB 538 GAGTACACGCCGATCCACTCAGCTGAACCTGGACCGGAGAGAGACGATCCCAATCAT 597  
QY 771 ACTCAGGAGAGAAATCTTTTACTGTGAGAGTGTGCTGACGACTTCTTGGAGGTGCA 830  
DB 598 ACTCAGGAGAGAAATCTTTTACTGTGAGAGTGTGCTGACGACTTCTTGGAGGTGCA 657  
QY 831 GCCAGATTCGGGGTGGCAGAAATCCGATTCATGTGCTCAGAAAGATCAAGGCGCTG 890  
DB 658 GCCAGATTCGGGGTGGCAGAAATCCGATTCATGTGCTCAGAAAGATCAAGGCGCTG 717  
QY 891 TCCCTGTGTTTAAATGCTGACACCACTTACTGTCAATGACCCGGGAATGACTGGGC 950  
DB 718 TCCCTGTGTTTAAATGCTGACACCACTTACTGTCAATGACCCGGGAATGACTGGGC 777  
QY 951 CAATCAGTGAAGTTGTGTGATTCGACACAGACATTTGGAGCTGCTTGAAGAAACAGAT 1010  
DB 778 CAATCAGTGAAGTTGTGTGATTCGACACAGACATTTGGAGCTGCTTGAAGAAACAGAT 837  
QY 1011 AATGATAGTGTGTAAGTGTGTTCTTTCTGTGTAAGTTCTGTCTGTGCAAGGCGAGTT 1070  
DB 828 AATGATAGTGTGTAAGTGTGTTCTTTCTGTGTAAGTTCTGTCTGTGCAAGGCGAGTT 897  
QY 1071 GATCAGTGAAGTCAAGAGAGAGCTTCTGTTTCTTAAGTGGCTGAGGGGCACTCTTA 1130  
DB 898 GATCAGTGAAGTCAAGAGAGAGCTTCTGTTTCTTAAGTGGCTGAGGGGCACTCTTA 957  
QY 1131 CTGGTAGGAAGATGACACAGAGAGCGCCTAGTGCAGAGAGTTGTGAAGACAGAGC 1190  
DB 958 CTGGTAGGAAGATGACACAGAGAGCGCCTAGTGCAGAGAGTTGTGAAGACAGAGC 1017  
QY 1191 AATGCAATGTGAAATTTGAGCGTTTCTTTCTTCTGATGTTCTCATGTTTGTGATG 1250  
DB 1018 AATGCAATGTGAAATTTGAGCGTTTCTTTCTTCTGATGTTCTCATGTTTGTGATG 1077  
QY 1251 TATATTAGTGAATTTACAGAGCTTACCTTGTGTAATTAAGTTACACCGTTGTGTT 1310  
DB 1078 TATATTAGTGAATTTACAGAGCTTACCTTGTGTAATTAAGTTACACCGTTGTGTT 1137  
QY 1311 TACATCTTTTGGAGGCCAGAGAAACGTTTGAAGAAACGATCACCTTCCAGATTCTCG 1370  
DB 1138 TACATCTTTTGGAGGCCAGAGAAACGTTTGAAGAAACGATCACCTTCCAGATTCTCG 1197  
QY 1371 GATTCGACCTTTTGCACACAGACATTTGCTTGGCAATCTTCCGGAATGATCATCAGC 1430  
DB 1198 GATTCGACCTTTTGCACACAGACATTTGCTTGGCAATCTTCCGGAATGATCATCAGC 1257  
QY 1431 AGCATCCCAACCGTGCACAGTGAATGCTTGTGCAAGAAAGATGATCTGAAATT 1490  
DB 1258 AGCATCCCAACCGTGCACAGTGAATGCTTGTGCAAGAAAGATGATCTGAAATT 1317  
QY 1491 CCTGTGTAGAAATTTAGCTTATCAATTCAGAGATAGCAAGTTTCACTGCCAATTTAGT 1550  
DB 1318 CCTGTGTAGAAATTTAGCTTATCAATTCAGAGATAGCAAGTTTAGTGCACAACTTTAGT 1377  
QY 1551 GGGTGAAGAAATTTAGTGTGTTTGGAGTGGAGTCACTTCTGTGTGTTCTTTTA 1610  
DB 1378 GGGTGAAGAAATTTAGTGTGTTTGGAGTGGAGTCACTTCTGTGTGTTCTTTTA 1437  
QY 1611 TGGGTGTTTGTATATCATGATCATAGCCAAATTTTGTGAAACGTTGTGTGAGA 1670  
DB 1438 TGGGTGTTTGTATATCATGATCATAGCCAAATTTTGTGAAACGTTGTGTGAGA 1497  
QY 1671 TAGTTGTTCTTTTACCCAGAGAGACATCAAGATACATTTGTAAATTAAGCTGATAGCA 1730  
DB 1498 TAGTTGTTCTTTTACCCAGAGAGACATCAAGATACATTTGTAAATTAAGCTGATAGCA 1557  
QY 1731 TATATTCAATCCTGTTGTACACTTGGTGAAGAAAGTATGAGCATGGAAGCTAAGATGAT 1790  
DB 1558 TATATTCAATCCTGTTGTACACTTGGTGAAGAAAGTATGAGCATGGAAGCTAAGATGAT 1617  
QY 1791 TAACTTAAGCTGTAATCATATGTTGTAGAGAAAGCTGTCTCCATGTCTAACAGAGACTTGA 1850

Db	1618	TAACTACCTGTCGTAATCATATGTTGTAGGAAAGCTGTTCCCATGCTCTACAGACTTGA	1677
Qy	1851	ATTCAAAGCATGTCAAGTGGATAGTAGATCTGTGGCGATATGAGAGGATCAGTGCCTT	1910
Db	1678	ATTCAAAGCATGTCAAGTGGATAGTAGATCTGTGGCGATATGAGAGGATCAGTGCCTT	1737
Qy	1911	TCCCATTCATTCTCGATGGAATTGTATCTACTAGTTTAACTATGTTAAATTTTTTCTAGT	1970
Db	1738	TCGGCATTCATTCTCGATGGAATTGTATCTACTAGTTTAACTATGTTAAATTTTTTCTAGT	1797
Qy	1971	TGTAATGTGPTATCTGGTAAATAGGTATATATATTTGGCCCTTACAATACCGTAACAATG	2030
Db	1798	TGTAATGTGPTATCTGGTAAATAGGTATATATATTTGGCCCTTACAATACCGTAACAATG	1857
Qy	2031	TTTGTCAATTTGAAATACHTTAATGCCAAGTAACAATGCATGCTTTGGAAATTTGGAAGAT	2090
Db	1858	TTTGTCAATTTGAAATACHTTAATGCCAAGTAACAATGCATGCTTTGGAAATTTGGAACAT	1917
Qy	2091	GGTTTATTTCTTTGAGAAGCAATATGTTTGCAATTAATGCTTTGATTTGTTTCATATCAAG	2150
Db	1918	GGTTTATTTCTTTGAGAAGCAATATGTTTGCAATTAATGCTTTGATTTGTTTCGATCAAG	1977
Qy	2151	AAATTGATTGAAGCTTCTCAAAACCTGTTTACGGTACTTGTGAAGAGGAGCCGTTTGG	2210
Db	1978	AAATTGATTGAAGCTTCTCAAAACCTGTTTACGGTACTTGTGAAGAGGAGCCGTTTGG	2037
Qy	2211	GAGAGACCAATGCGATCGCT- GTCCAAAGTGTTCCTTTTAAGTGCCTTTTAACTGGAGAGG	2269
Db	2038	GAGAGACCAATGCGATCGCTCNCNCAAGTGTTCTTTGTTAAAGTTCCTTTTAACTGGAGAGG	2097
Qy	2270	CTAAACCTCAAAATACCTTTTTTTTAACTGCATCTATATAAATGGGCACAGTAGTGCCTT	2329
Db	2098	CTAAACCTCAAAATACCTTTTTTTTAACTGCATCTATATAAATGGGCACAGTAGTGCCTT	2157
Qy	2330	AC 2331	
Db	2158	AC 2159	

```

RESULT 2
US-09-602-877A-83/C
; Sequence 83, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF BREAST CANCER
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(251)
; OTHER INFORMATION: n = A,T,C or G
US-09-602-877A-83

```

QY	2141	TCATATCAAGAAATTGATTGAACGTTTCTCAAAACCCCTGTTTACGGTACTTTGGTAAGAGCGGA	2200
Db	191	TCATATCAAGAAATTGATTGAACGTTTCTCAAAACCCCTGTTTACGGTACTTTGGTAAGAGCGGA	132
QY	2201	GCCGGTTGGGAGAGACCAATTGCATCGCTGTCCAAGTGTTTCTGTTAAAGTGCCTTTTAAA	2260
Db	131	GCCGGTTGGGAGAGACCAATTGCATCGCTGTCCAAGTGTTTCTGTTAAAGTGCCTTTTAAA	72
QY	2261	CTGAGAGGCTTAACCTCAAAATACTTTTTTAACTGCATTCTATATAAATATGGGCACAGT	2320
Db	71	CTGAGAGGCTTAACCTCAAAATATTTTTTTTAACTGCATTCTATATAAATATGGGCACAGT	12
QY	2321	ATGCTCCTTAC	2331
Db	11	ATGCTCCTTAC	1

RESULT 3  
 US-08-232-463-14/c  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEFFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299

Query Match	10.6%	Score 246.8;	DB 4;	Length 251;
Best Local Similarity	98.8%;	Pred. No. 4.4e-66;		
Matches 248;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
2081	TTTGGAGATGGTTTTATTCTTTGAGAGCAAAATATGTTTGCAATTAATGCTTTCATTGT	2140		
251	TTTGGAAAATGGTTTTATTNTTTTGAGAGCAAAATATGTTTGCAATTAATGCTTTCATTGT	192		

```

QY 545 CATCTCCAGCTGGGCGGGAAGATGATTCAGCAGCAGTACACCACTCC 604
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1365 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1306
QY 605 CAGCGTGTGTCATGTATGTAGAGTATCAGAGAGAGAGAGAGAGAGAGAG 664
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1305 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1246
QY 665 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1245 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1186
QY 725 CACCTCAGCTGAGTACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 784
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1185 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1126
QY 785 TCTTTTACTGTGAGAGTGTGTCACGACTTCTCGAGTGGCAGCCGAGATCGGGGT 844
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1125 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1066
QY 845 GCGAGAAATCCCATGTTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 904
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1065 TCGCAGAGCTCTCGACAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1006
QY 905 TGCTGACACCACTTACTGT 924
Db 1005 GAAGGAAATAGTTATAGT 986

```

## RESULT 4

```

US-08-939-309-7/C
Sequence 7, Application US/08939309
Patent No. 6423527
GENERAL INFORMATION:
APPLICANT: Saba, Julie D.
TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
TITLE OF INVENTION: POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/939,309
FILING DATE: 29-SEP-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: David, Mark J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 200116.402
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1770 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1767

```

## US-08-939-309-7

Query Match 1.8%; Score 41; DB 4; Length 1770;  
 Best Local Similarity 51.4%; Pred. No. 0.028;  
 Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```

QY 2011 CTTCACATACCGTAACTGTTGTCTTTGAAATACCTTAAGCAGTAATGCAT 2070
   |||||
Db 2332 CTTCACATACCGTAACTGTTGTCTTTGAAATACCTTAAGCAGTAATGCAT 173
QY 2071 GCTTGAAATTTGAGATGTTTATCTTTGAGAGCAATATGTTGATTAATG 2130
   |||||
Db 1772 CTTTCAACATGTTTACATGAGGTATTTCTGAGTACCATGATAGCTATTTTAA 113
QY 2131 CTTGATGTTTATATCAGAAATGATTAAGCTTCAACCTGTTTACGTAATG 2190
   |||||
Db 112 CGTTGATGTTTATATCAGAAATGATTAAGCTTCAACCTGTTTACGTAATG 53
QY 2191 GTTAA 2195
Db 52 GCATG 48

```

## RESULT 5

```

US-08-665-040-9/C
Sequence 9, Application US/08665040
Patent No. 5869318
GENERAL INFORMATION:
APPLICANT: ESTIVILL, PALLEJA, XAVIER
APPLICANT: FUENTES, JUAN JOSE
APPLICANT: PRITCHARD, MELANIE
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING PCR" TECHNIQUE,
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSGR1) HIGHLY
TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LADAS & PARRY
STREET: 26 WEST 61ST STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1 FOR DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,040
FILING DATE: JUNE 7, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: ES P9501140
FILING DATE: JUNE 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: JANET I. CORD
REGISTRATION NUMBER: 33,778
REFERENCE/DOCKET NUMBER: U010815-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULAR TYPE: DNA (genomic)
US-08-665-040-9

```

Query Match 1.7%; Score 40; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 0.005;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 GAGTCCCTTTAGGACGTATGACAAAGACATCACCTTTC 279  
 DB 40 GAGTCCCTTTAGGACGTATGACAAAGACATCACCTTTC 1

## RESULT 6

US-08-991-677-11/c  
 ; Sequence 11, Application US/08991677A  
 ; Patent No. 6252135

; GENERAL INFORMATION:  
 ; APPLICANT: Chiang, Vincent L  
 ; APPLICANT: Cartaway, Daniel T  
 ; APPLICANT: Smeltzer, Richard H  
 ; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms  
 ; FILE REFERENCE: 50617  
 ; CURRENT APPLICATION NUMBER: US/08/991,677A  
 ; CURRENT FILING DATE: 1997-12-16  
 ; EARLIER APPLICATION NUMBER: US 60/033,381  
 ; EARLIER FILING DATE: 1996-12-16  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 11

; LENGTH: 2251  
 ; TYPE: DNA  
 ; ORGANISM: Pinus taeda  
 US-08-991-677-11

Query Match 1.7%; Score 39.6; DB 4; Length 2251;  
 Best Local Similarity 53.2%; Pred. No. 0.088;  
 Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1952 TTGTGAATTTTCTAGTTGTAATGTCGTGTAATAGGTATATATTTGGCC 2011  
 DB 326 TGTTAGCTTATTTTAAAGTTTGTGAATTTGTATAAAGTTGTTATTTGAAG 267

QY 2012 TTACAATACGTACATGTTGTGTCATTTTGAATATCTTAATGCCAAGTAAACATG 2071  
 DB 266 TTTTAATAAATTTGTATAATATTTTTTTTAAATAGTTAAGTTTATTTATTAAATC 207

QY 2072 CTTTGGAAATTTGGAATGCTTTATCTTCTTGAGAAG 2109  
 DB 206 TATTTAATTTGTCGAGAGGTTGTGATGAGAAG 169

## RESULT 7

US-08-232-463-14  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367

; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEIFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,463  
 ; FILING DATE:

; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/935,313  
 ; FILING DATE:  
 ; APPLICATION NUMBER: EP 91 114 300.6  
 ; FILING DATE: 26-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)836-9300  
 ; TELEFAX: (703)683-4109  
 ; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 7218 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; CLONE: PT29pt-F1s  
 US-08-232-463-14

Query Match 1.6%; Score 38.2; DB 1; Length 7218;  
 Best Local Similarity 8.8%; Pred. No. 0.5;  
 Matches 25; Conservative 140; Mismatches 118; Indels 0; Gaps 0;

QY 1214 TTTCCTTTCTCCCTCATGTTCTCATGTTTGGCATGATATATTACTGATTACAGACTA 1273  
 DB 1205 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1264

QY 1274 ACCTTGTTCGTATATAAGTTACACCTGTTGTTTGTATCATCTTTGGGAGCCAGAA 1333  
 DB 1265 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1324

QY 1334 AGCGTTTGGAAACGATACCTTTCCAGATTCTCGGATTCTCGACTCTTTCGACACG 1393  
 DB 1325 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1384

QY 1394 ACTTCCTTCGGAAGTCTTCTCGAATCATCTAGCATCCCAACCTGCAACGTG 1453  
 DB 1385 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1444

QY 1454 TAACCTGTGCTTTTGGAAAGAGTGTGATCTGAAATTCCTCTG 1496  
 DB 1445 CTTCTATCTCTTAACTACTTGCATAGTAGTAAATTACAGTG 1487

## RESULT 8

US-09-889-595-7/c  
 ; Sequence 7, Application US/09889595  
 ; Patent No. 6410749

; GENERAL INFORMATION:  
 ; APPLICANT: Aventis CropScience GmbH  
 ; TITLE OF INVENTION: PROMOTERS FOR GENE EXPRESSION IN CARYOPSES OF PLANTS  
 ; FILE REFERENCE: 514413-3885  
 ; CURRENT APPLICATION NUMBER: US/09/889,595  
 ; CURRENT FILING DATE: 2001-07-05  
 ; PRIOR APPLICATION NUMBER: DE 100 32 379.0  
 ; PRIOR FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 7

; LENGTH: 327  
 ; TYPE: DNA  
 ; ORGANISM: Triticum aestivum  
 US-09-889-595-7

Query Match 1.6%; Score 38; DB 4; Length 327;  
 Best Local Similarity 47.5%; Pred. No. 0.079;  
 Matches 113; Conservative 0; Mismatches 125; Indels 0; Gaps 0;





```

REFERENCE/DOCKET NUMBER: 960296.95238
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 3770..4013
OTHER INFORMATION: /note="RP4 origin of DNA transfer (orit) from
US-09-056-075-1
Query Match
Best Local Similarity 56.6%; Pred. No. 0.93; Length 6243;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 2038 TTTTGAATATCTTAATGCAAGTACGATGCTTGAATTTGGAAGATGCTTTTA 2097
Db 1499 TTTTCATATACGTAATATGATGCTTAAATGTTTATGTTACTTCAATAGAGTAA 1558

Qy 2098 TTTTTCAGACGAATATGTTGCAATGCTTGAATGTTTCATTCAGAAATGA 2157
Db 1559 TTTCTTATTTTAAAGCCTTTTCTTAAAGGCTTTATTTTCTTAAATACATTTA 1618

Qy 2158 TT 2159
Db 1619 AT 1620

RESULT 14
US-09-007-005-17
Sequence 17, Application US/0907005B
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihc
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
EARLIER FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match
Best Local Similarity 1.6%; Score 36.8; DB 4; Length 289;
Matches 15; Conservative 101; Mismatches 122; Indels 0; Gaps 0;

Qy 615 GTCCATGATATGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
Db 27 RURURARCRARARURURARCRARARURGRNRNRNRNRNRNRNRNRNRNR 86

```

```

Qy 675 AGACCTAAGCCAAAATATTCAGACGAGAGCCGAGTACGCCGATTCACCTCAGC 734
Db 87 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 146
Qy 735 TGAATCGCAGCAGCAGAGAGCCATTCCTCAATCATCTACGAGGAGATCTTTACTG 794
Db 147 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 206
Qy 795 TGAGGTGCTGTCACGACTCTTCGAGGTGCGACCCGAGATCGGGGTGAGAGAA 852
Db 207 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 264

RESULT 15
US-09-007-005-17/c
Sequence 17, Application US/0907005B
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihc
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
EARLIER FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match
Best Local Similarity 10.4%; Score 36.8; DB 4; Length 289;
Matches 26; Conservative 94; Mismatches 130; Indels 0; Gaps 0;

Qy 1277 TTTGTCGATATTAAGTACCGCTTGTGTTTACATCTTTGGGAAGCAGAGAAC 1336
Db 265 TTTTAVGCGYCYGAYAGVGYTYAVCYGAYGYTYGYSYNNVSYNNVSYNNV 206
Qy 1337 GTTTGAAAAGCATCACCTTCCAGATTCGAGATTCGACTCTTGCACACGACT 1396
Db 205 YNNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNV 146
Qy 1397 TGCTTGGCACTCTTCGAGATTCATCTACGATCCCAACCGTGCAACGTGTA 1456
Db 145 YNNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNV 86
Qy 1457 CTGTGCTTTTGGAAAAGAGTGAATTCCTCTGTGAGATTTAGCTTATACAT 1516
Db 85 YNNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNV 26
Qy 1517 TCAGAGATA 1526
Db 25 YAGTYAYTA 16

Search completed: December 14, 2002, 18:14:52
Job time : 123.038 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 18.5118 Seconds  
(without alignments)  
9923.385 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gaggtgcaaggaacctcca.....gaggaggaaggaagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PTCTUS COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340.2	56.8	2174	2	US-08-665-040-1
2	36.6	6.1	3303	1	US-08-081-610-3
3	34	5.7	2118	4	US-09-221-017B-789
4	33	5.5	1685	1	US-09-640-305-1
5	33	5.5	1685	1	US-08-360-673-1
6	32.4	5.4	388	4	US-09-325-932A-108
7	32.2	5.4	3177	4	US-09-676-974-1
8	32	5.3	815	4	US-09-383-586-9
9	32	5.3	36519	3	US-08-923-137-2
10	31.8	5.3	4214	4	US-09-221-017B-293
11	31.6	5.3	1279	3	US-08-961-083-47
12	31.6	5.3	16995	4	US-08-961-527-82
13	31.2	5.2	4221	4	US-09-651-658-22
14	31.2	5.2	4221	4	US-09-650-855-22
15	31.2	5.2	16389	4	US-09-741-154-3
16	31	5.2	1930	2	US-08-737-716-1
17	31	5.2	1930	2	US-08-737-716-11
18	30.4	5.1	7431	4	US-08-665-040-9
19	30.4	5.1	7431	4	US-09-306-998-2
20	30.2	5.0	2518	4	US-09-433-699-3
21	30	5.0	1284	3	US-09-196-857-1
22	30	5.0	3489	2	US-08-728-323A-1
23	30	5.0	3489	4	US-09-298-568-1
24	30	5.0	32207	2	US-08-770-379-20
25	30	5.0	32207	4	US-08-757-669A-20
26	30	5.0	32207	4	US-09-230-371A-20
27	29.8	5.0	1512	2	US-08-909-965C-8

28	29.8	5.0	4031	2	US-08-993-118-1	Sequence 1, Appli
29	29.8	5.0	4031	3	US-08-845-528C-1	Sequence 1, Appli
30	29.8	5.0	4225	2	US-08-993-118-9	Sequence 9, Appli
31	29.8	5.0	4225	3	US-08-845-528C-9	Sequence 9, Appli
32	29.8	5.0	4265	4	US-09-061-709-1	Sequence 1, Appli
33	29.8	5.0	4415	4	US-09-486-580A-1	Sequence 1, Appli
34	29.8	5.0	12286	4	US-09-322-478-17	Sequence 17, Appli
35	29.8	5.0	12571	4	US-09-322-478-20	Sequence 20, Appli
36	29.6	4.9	2082	4	US-09-440-325A-2	Sequence 2, Appli
37	29.4	4.9	424	4	US-09-397-787-133	Sequence 133, App
38	29.4	4.9	1829	1	US-07-688-352C-15	Sequence 15, Appli
39	29.4	4.9	1829	2	US-08-474-379C-15	Sequence 15, Appli
40	29.4	4.9	1829	3	US-09-146-249A-15	Sequence 15, Appli
41	29.4	4.9	1829	3	US-08-206-188B-15	Sequence 15, Appli
42	29.4	4.9	1829	5	PCT-US91-02714-15	Sequence 15, Appli
43	29.4	4.9	2269	4	US-09-394-645-1	Sequence 1, Appli
44	29.4	4.9	2269	4	US-09-243-560B-1	Sequence 1, Appli
45	29.2	4.9	588	4	US-09-385-982-286	Sequence 286, App

## ALIGNMENTS

### RESULT 1

US-08-665-040-1  
; Sequence 1, Application US/08665040  
; Patent No. 5969318  
; GENERAL INFORMATION:  
; APPLICANT: ESTIVILL PALLEJA, XAVIER  
; APPLICANT: FUENTES, JUAN JOSE  
; APPLICANT: PRITCHARD, MELANIE  
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING PCR" TECHNIQUE,  
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
; TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD  
; TITLE OF INVENTION: FOR CHARACTERIZING IT.  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY  
; STREET: 26 WEST 61ST STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10023  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT 5.1 FOR DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,040  
; FILING DATE: JUNE 7, 1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: ES P9501140  
; FILING DATE: JUNE 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JANET I. CORD  
; REGISTRATION NUMBER: 33,778  
; REFERENCE/DOCKET NUMBER: U010815-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 708-1800  
; TELEFAX: (212) 246-8959  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2174 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: cDNA for mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO

```

ORIGINAL SOURCE: human
ORGANISM: Homo sapiens
DEVELOPMENTAL STAGE: foetal
TISSUE TYPE: Brain
IMMEDIATE SOURCE: gene library of cDNA
LIBRARY: gene library of cDNA from foetal
CLONE: BC-17.8-1 and BC-17.8-2
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
MAP POSITION: 21q22.1-q22.2
FEATURE:
NAME/KEY: cDNA for mRNA, BC-17.8
LOCATION: 1..2174
FEATURE:
NAME/KEY: untranslated 5'
LOCATION: 1..48
FEATURE:
NAME/KEY: coding sequence
LOCATION: 49..560
OTHER INFORMATION: Down Syndrome critical
region 1 (DSCR1)
FEATURE:
NAME/KEY: DSCR1
LOCATION: 1..171 PEPTIDES
IDENTIFICATION METHOD: translation of the
OTHER INFORMATION: - deduced protein
OTHER INFORMATION: - proline-rich protein domains
OTHER INFORMATION: - glutamic acid-rich protein domains
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
FEATURE:
NAME/KEY: untranslated 3'
LOCATION: 564..2174
FEATURE:
NAME/KEY: 2 poly (A)
LOCATION: 1541..1546 AND 2132..2137
US-08-665-040-1

```

```

Query Match          56.8%; Score 340.2; DB 2; Length 2174;
Best Local Similarity 88.6%; Pred. No. 5..4e-107;
Matches 380; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```

```

QY 167 GCCAAATTGATTCCTCTTGAACAATGACAGACACACCTTCCAGTATTTTAAAG 226
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 58 GCCAAATTGATTCCTCTTGAACAATGACAGACACACCTTCCAGTATTTTAAAG 117
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 227 AGCTTCAAGCTGTCGGGATAAATCTCAGCAACCCCTTATCTGAGCGGATGCGAGCTG 286
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 118 AGCTTCAAGCTGTCGGGATAAATCTCAGCAACCCCTTATCTGAGCGGATGCGAGCTG 177
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 287 CGGCTGCAAGACCGAGTTCCTGAGGAAGGAATGAAGTTGTTTCTCAGACTTGA 346
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 178 CAGCTGATATGAGTGAATTTCTGGGAAAGGAATGAAGTTGTTTCTCAGACTTGA 237
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 347 CACATAGGAAGTTGACACTTGGCTCCGCC -AATCCGACAAAGAGTTCCATCTCCCT 405
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 238 CACATAGGAAGTTGACACTTGGCTCCGCCCAATCAGACAGAGCTTTCATCTCCCT 297
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 406 CCGGCTCTCTCCCGTGGCTGGAAACAAGTAGAATGACACCCCGTCATTAATAC 465
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 288 CCGGCTCTCTCCCGTGGCTGGAAACAAGTAGAATGACACCCCGTCATTAATAC 357
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 466 GATCTTTATATGCAATCTCAAGTGGGCGCAGAGAGAAATGAACTGATGAGCGG 525
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 358 GATCTTTATATGCAATCTCAAGTGGGCGCAGAGAGAAATGAACTGATGAGCGG 417
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 526 ACAGACACACTCCCAAGTGTGTGTGTCACCGTGTGAGAGTACCAAGAGATGAGAG 585
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 418 ACTGACACCACTCCCAAGTGTGTGTGTCACCGTGTGAGAGTATGAGATGAGAGAGAA 477
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 586 GAAAGAGAA 594
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 478 GAAAGAGAA 486
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 2
US-08-081-610-3
Sequence 3, Application US/08081610
Patent No. 5445941
GENERAL INFORMATION:
APPLICANT: Yang, Na N
TITLE OF INVENTION: Materials and Methods for Screening
TITLE OF INVENTION: Anti-Osteoporosis or Serum Lipid Lowering Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Allegretti and Witcoff, Ltd.
STREET: 10 S. Wacker Dr.
CITY: Chicago
STATE: IL
COUNTRY: U.S.A
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,610
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heaphy, Barbara A
REGISTRATION NUMBER: 34,619
REFERENCE/DOCKET NUMBER: 93,402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: mRNA
LOCATION: 2170..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2214..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2219..3303
FEATURE:
NAME/KEY: misc RNA
LOCATION: 3301..3303
OTHER INFORMATION: /note= "CDS Start, codon start = 1,
OTHER INFORMATION: translation M"
FEATURE:
NAME/KEY: TATA signal
LOCATION: 2170..2176
FEATURE:
NAME/KEY: misc feature
LOCATION: 1896..2306
OTHER INFORMATION: /note= "PB-301 -301 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 1976..2306
OTHER INFORMATION: /note= "PB-221 -221 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2106..2306
OTHER INFORMATION: /note= "PB-91 -91 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2137..2306
OTHER INFORMATION: /note= "PB-60 -60 to +110"

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2150..2306
; OTHER INFORMATION: /note= "pB-47 -47 to +110"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2159..2306
; OTHER INFORMATION: /note= "pB-38 -38 to +110"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2159..2271
; OTHER INFORMATION: /note= "TGPB-3 position -38 to
; OTHER INFORMATION: +75"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2159..2231
; OTHER INFORMATION: /note= "TGPB-3 position -38 to
; OTHER INFORMATION: +35"
US-08-081-610-3

```

```

Query Match 6.1%; Score 36.6; DB 1; Length 3303;
Best Local Similarity 60.6%; Pred. No. 0.032;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 499 GCAGAGAAGTAACTGCATGCGAGCGACACACACCACTCCAGTGTGGTGTCCACGTG 558
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2161 GAGGGAGGTATAAAATTCAGCAGAGAGAGAAATAGAGAAAGCAGTGTGTGCATGTG 2220
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 559 TGTGAGAGTACCAAGAGAATGAGGAGCAAGAGGAAGAG 597
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2221 TGTGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 2259
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 3
US-09-221-017B-789/c
; Sequence 789, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Moroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 789:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2118
US-09-221-017B-789

Query Match 5.7%; Score 34; DB 4; Length 2118;
Best Local Similarity 57.5%; Pred. No. 0.19;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 226 GAGCTTCAAACTGTCGCGATAAACTTCAGCAACCCCTTATCTGCAGCGGATGCCAGGCT 285
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1380 GAGTCTCTATATGTCGAGGAGGAAAATCCGTACCACGATGTAGACAATACAGGGGT 1321
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 286 GCGGTGCACAAGACCGAGTTCCTGGGAAGGAATGAAGTTGTAT 331
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1320 TCGTCGCAAGTGCAGGAAATCCTTCCGAGAAAATCAACTTTTTT 1275
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-640-305-1/c
; Sequence 1, Application US/09640305
; Patent No. RE37447
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/640,305
; FILING DATE: 16-AUG-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808

```

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1685 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Kluyveromyces fragilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1683  
OTHER INFORMATION: /product= "Protease B gene"  
/gene= "K1.PR81"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-640-305-1

Query Match 5.5%; Score 33; DB 1; Length 1685;  
Best Local Similarity 53.5%; Pred. No. 0.37;  
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 88 TAGGAGCTTAGTACATTTTATGCTCCCTGATGCTTGTGTGGCAACGATGCTT 147  
DB 595 TAGAATCTTCTTCAACAAATTCACCAATGGGCTTCTGTGTAATGATGATCTT 536  
QY 148 CAGCGAAGTGAGACGAGGCCAAATTTGATCCCTCTTCAAGACATATGACAGACAC 207  
DB 535 TTGTGAAGTAGCCACATTAACGAAAGGACCGCAATATCATAAGAGTCTTGAATAC 476  
QY 208 CACCTTCCA 216  
DB 475 CACCTTCCA 467

RESULT 5  
US-08-360-673-1/c  
Sequence 1, Application US/08360673  
Patent No. 5679544

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard  
APPLICANT: Fournier, Alain  
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR  
TITLE OF INVENTION: PREPARATION AND USE  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Rd. 3c43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19002

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/360.673  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR93/00623  
FILING DATE: 23-JUN-1993  
PRIOR APPLICATION DATA: FR 92/07785  
APPLICATION NUMBER:  
FILING DATE: 25-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Julie K.  
REGISTRATION NUMBER: 38,619  
REFERENCE/DOCKET NUMBER: ST92040-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839

TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1685 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Kluyveromyces fragilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1683  
OTHER INFORMATION: /product= "Protease B gene"  
/gene= "K1.PR81"  
US-08-360-673-1

Query Match 5.5%; Score 33; DB 1; Length 1685;  
Best Local Similarity 53.5%; Pred. No. 0.37;  
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 88 TAGGAGCTTAGTACATTTTATGCTCCCTGATGCTTGTGTGGCAACGATGCTT 147  
DB 595 TAGAATCTTCTTCAACAAATTCACCAATGGGCTTCTGTGTAATGATGATCTT 536  
QY 148 CAGCGAAGTGAGACGAGGCCAAATTTGATCCCTCTTCAAGACATATGACAGACAC 207  
DB 535 TTGTGAAGTAGCCACATTAACGAAAGGACCGCAATATCATAAGAGTCTTGAATAC 476  
QY 208 CACCTTCCA 216  
DB 475 CACCTTCCA 467

RESULT 6  
US-09-325-932A-108/c  
Sequence 108, Application US/09325932A  
Patent No. 6451604

GENERAL INFORMATION:

APPLICANT: Flinn, Barry  
APPLICANT: Lasham, Annette  
TITLE OF INVENTION: Compositions affecting programmed cell  
TITLE OF INVENTION: death and their use in the modification of forestry plant develo  
FILE REFERENCE: 1022  
CURRENT APPLICATION NUMBER: US/09/325.932A  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 206  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 108  
LENGTH: 388  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-325-932A-108

Query Match 5.4%; Score 32.4; DB 4; Length 388;  
Best Local Similarity 50.6%; Pred. No. 0.25;  
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 364 CCTGGCTCCGCAATCCGACAAAGTCTCATCTCCCTCGGCTTCTCCGCTT 423  
DB 379 CCGAGCTTGGCAGGTAGATGATCTGCTTCACTAAGATCCGCAATATCTATGAT 320  
QY 424 GGCTGGAACAAAGTAGAATGCCACCCCTGATTAATTTGATGATGATGATGATG 483  
DB 319 AATTTGAAAGGCGGTGTCCATTAAGCCCTGTTGCAACTGCGTCTTTACTGCAATC 260  
QY 484 TCCAGCTGGGCGGACGAGAGAGATGATGATGATGATGATGATGATGATGATG 517  
DB 259 TACCACTGTGCTGTGATTAAGAAACGATGATGATGATGATGATGATGATGATG 226

RESULT 7  
US-09-676-974-1/c

; Sequence 1, Application US/09676974  
; Patent No. 6391631  
; GENERAL INFORMATION:  
; APPLICANT: HACKER, JORG  
; APPLICANT: SONNEN-BORN, ULRICH  
; APPLICANT: SCHULZE, JURGEN  
; APPLICANT: BLUM-OEHLER, GABRIELE  
; APPLICANT: MALINKA, JURGEN  
; APPLICANT: PROPERT, HANS  
; TITLE OF INVENTION: BACTERIAL PLASMIDS  
; FILE REFERENCE: 11347/268416/BET  
; CURRENT APPLICATION NUMBER: US/09/676,974  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP98/01720  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: DE 197 13543.9  
; PRIOR FILING DATE: 1997-04-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3177  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; US-09-676-974-1

Query Match 5.4%; Score 32.2; DB 4; Length 3177;  
Best Local Similarity 59.1%; Pred. No. 1;  
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Qy 69 TGAAGAGCAGAGATGATTTAGGACCTTAGCTACAAATTTAGCTCCCTGATGCTTG 128  
Db 2938 TCAAAATGCTCAATGAATCATCGATTTATCGAGTCTTTTCAGAAATATCTGGATG 2879  
Qy 129 TGGCAACGATGATGCTTTCAGCGAAAGTGAGA 161  
Db 2878 TAATAAGGTTGAAGGTCTTCAGAGAATGAGACA 2846

RESULT 8  
US-09-383-586-9  
; Sequence 9, Application US/09383586  
; Patent No. 6242419  
; GENERAL INFORMATION:  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Abernethy, Nevin  
; APPLICANT: Onrust, Rene  
; APPLICANT: Kumble, Anand  
; APPLICANT: Murison, Greg  
; TITLE OF INVENTION: Compounds isolated from stromal cells  
; FILE REFERENCE: 11000.1037c1  
; CURRENT APPLICATION NUMBER: US/09/383,586  
; CURRENT FILING DATE: 1999-08-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 815  
; TYPE: DNA  
; ORGANISM: Mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (119)...(682)  
; NAME/KEY: sig peptide  
; LOCATION: (119)...(205)  
; US-09-383-586-9

Query Match 5.3%; Score 32; DB 4; Length 815;  
Best Local Similarity 54.2%; Pred. No. 0.53;  
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
Qy 141 ATGCTTACCGAAGTGAAGACAGCGGCCAAATTTGAATCCCTCTTCAGAACATATGACA 200

Db 431 AAGTCTACACCCAAATGCAAGTGCTTCTCCAAACTCTACCCACAGCTCAGCATCCATGACA 490  
Qy 201 AGCAGACACCTTCCAGTATTTTATAGAGCTTCAACGCTGTCGGGATAAACTTCAGCAACC 260  
Db 491 ACCCAGCCCCACAGTAGTATTATTGACAACTGTAACGGTTTCAGCAACTACTCATCCACC 550

## RESULT 9

US-08-923-137-2  
; Sequence 2, Application US/08923137  
; Patent No. 6083716  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Farina, Steven P.  
; APPLICANT: Fisher, Krishna J.  
; TITLE OF INVENTION: Chimpanzee Adenovirus Vectors  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr., P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: United States of America  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923.137  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/024,700  
; FILING DATE: 06-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: GNVPN.021C1PIUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; US-08-923-137-2

Query Match 5.3%; Score 32; DB 3; Length 36519;  
Best Local Similarity 56.7%; Pred. No. 5.4;  
Matches 59; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
Qy 494 GCCCAGGAGAGAGTATGAACTGCATGCAGCGACAGACACCACTCCAGCTGTGTGTGCC 553  
Db 25600 GTCAGGAGAGGAGGAGGAGATGCGAGGAAGTGGGAGCACTTCAGGAGAGGAGGACA 25659  
Qy 554 ACCTGTGTGAGAGTGACCAAGAGATGAGGAGGAAGAGAGAG 597  
Db 25660 GCCTGCAAGACAGTCTGGAGGAGACGAGGAGGAGGAGGAG 25703

## RESULT 10

US-09-221-017B-293/c  
; Sequence 293, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: ROSS, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221.017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P22911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 293:
SEQUENCE CHARACTERISTICS:
LENGTH: 4214 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
TOPOLOGY: circular
ANTHROPATHIC: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...4214
US-09-221-017B-293

Query Match 5.3%; Score 31.8; DB 4; Length 4214;
Best Local Similarity 59.3%; Pred. No. 1.7; 37; Indels 0; Gaps 0;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 30 TGAAGTGAAGAGCGAGTCTGCTTAAAGCGTCTGCGGCGGTAAGAAAGCAGAAATGATTTTA 89
DB 3420 TCGCTGGAAGAGAAAGATGATGCTGAGCAAGCTGCTCCCATATTCAGGCGACGTGATTCGA 3361
QY 90 GGGACCTTACCTACATTTTAAAGCTCCCTGAT 120
DB 3360 TGAAGCTTACCTCCCATGCTTCTCTGAAGCAT 3330

RESULT 11
US-08-961-083-47/c
Sequence 47, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brooks, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 1279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-47

Query Match 5.3%; Score 31.6; DB 3; Length 1279;
Best Local Similarity 49.4%; Pred. No. 0.96;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 53 TTAAGCGTCTGCGCCGCTGAAAGAGAGATTTTAAAGGACTTTAGTACATTTTACG 112
DB 733 TTTAGCTTTTACTTCTGAGATGATGTCACGAAATTTAGCTTTTACCTGCAAGCTTTC 674
QY 113 TCCCGATTTGCTTGTGCGAAAGATGATCTTCAGCGAAAGTGAAGCAGGCGCAAA 172
DB 673 TCGTTGATGCTTTCTGGAATGTTACATTAACATCAACGATTTGCGCAGCTGATGACC 614
QY 173 TTTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCAGT 218
DB 613 TACCAATGCTCTTCGAAACCGAGGATGATGACCTGAAACCAAGT 568

RESULT 12
US-08-961-527-82/c
Sequence 82, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 18.45 Seconds  
(without alignments)  
9923.385 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggaatttaaggacttag.....cacgcacccacttagctga 597

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	398	66.7	2174	2	US-08-665-040-1
2	38.4	6.4	3303	1	US-08-081-610-3
3	37.8	6.3	7218	1	US-08-232-463-14
4	37.6	6.3	7130	4	US-09-056-105-31
5	36.2	6.1	6328	4	US-08-913-832A-1
6	36.2	6.1	6328	4	US-09-249-181A-1
7	34	5.7	1296	1	US-07-816-283-9
8	34	5.7	1296	1	US-08-417-103-9
9	34	5.7	1218	4	US-09-221-017B-789
10	33.2	5.6	2277	1	US-08-676-967-2
11	33.2	5.6	2277	1	US-08-676-974-2
12	33.2	5.6	2277	2	US-09-098-487-2
13	33.2	5.6	2289	4	US-09-312-038-3
14	33	5.5	1685	1	US-09-840-305-1
15	33	5.5	1685	1	US-08-360-673-1
16	33	5.5	2089	1	US-08-552-142A-1
17	33	5.5	2089	1	US-08-910-973-1
18	33	5.5	2089	4	US-09-499-227-1
19	33	5.5	2089	5	PCT-US95-05741-1
20	33	5.5	4221	4	US-09-651-656-22
21	33	5.5	4221	4	US-09-650-855-22
22	32.4	5.4	435	4	US-08-961-527-20
23	32.2	5.4	435	4	US-09-397-787-282
24	32	5.4	815	4	US-09-383-586-9
25	32	5.4	16442	3	US-08-781-891-208
26	31.8	5.3	3489	2	US-08-728-323A-1
27	31.8	5.3	3489	4	US-09-298-568-1

Sequence 20, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 11, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 2206, Ap  
Sequence 7, Appl  
Sequence 11, Appl  
Sequence 209, App  
Sequence 9, Appl  
Sequence 17, Appl

Sequence 20, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 11, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 2206, Ap  
Sequence 7, Appl  
Sequence 11, Appl  
Sequence 209, App  
Sequence 9, Appl  
Sequence 17, Appl

Sequence 20, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 11, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 1, Appl  
Sequence 2206, Ap  
Sequence 7, Appl  
Sequence 11, Appl  
Sequence 209, App  
Sequence 9, Appl  
Sequence 17, Appl

#### ALIGNMENTS

RESULT 1  
US-08-665-040-1  
; Sequence 1, Application US/08665040  
; Patent No. 5869318  
; GENERAL INFORMATION:  
; APPLICANT: ESTIVILL PALLEJA, XAVIER  
; APPLICANT: FUENTES, JUAN JOSE  
; APPLICANT: PRITCHARD, MELANIE  
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPlicing PCR" TECHNIQUE,  
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LADAS & PARRY  
; STREET: 26 WEST 61ST STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10023  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT 5.1 FOR DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,040  
; FILING DATE: JUNE 7, 1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: ES P9501140  
; FILING DATE: JUNE 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JANET I. CORD  
; REGISTRATION NUMBER: 33,778  
; REFERENCE/DOCKET NUMBER: U010815-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 708-1800  
; TELEFAX: (212) 246-8959  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2174 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: Double  
; TOPOLOGY: circular  
; MOLECULE TYPE: cDNA for mRNA  
; HYPOTHETICAL: No  
; ANTI-SENSE: No



```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2137..2306
; OTHER INFORMATION: /note= "pb-60 -60 to +110"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2150..2306
; OTHER INFORMATION: /note= "pb-47 -47 to +110"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2159..2306
; OTHER INFORMATION: /note= "pb-38 -38 to +110"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2159..2271
; OTHER INFORMATION: /note= "TGFB-3 position -38 to
; OTHER INFORMATION: +75"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2159..2231
; OTHER INFORMATION: /note= "TGFB-3 position -38 to
; OTHER INFORMATION: +35"
US-08-081-610-3

```

```

Query Match 6.4%; Score 38.4; DB 1; Length 3303;
Best Local Similarity 58.9%; Pred. No. 0.015; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 46;

```

```

Qy 421 GGAGAGAAGTATGAATCTGACGCGAGACAGACCCCACTCCAGTGTGGTCCACGTG 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2161 GGAGGAGGTATAAATTCAGCAGAGAGAAATAGAGAAACAGTGTGTGTCATGTG 2220
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 481 TGTGAGTGTACCAAGAGAAATGAGCGAGGAGAGAGAGAGATGGAGAGAAATGA 532
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2221 TGTGTGAGAGAGAGGAGGAGAGCGAGGAGGAGGAGGAGGAGGAGAGAGA 2272
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (703)836-9300
; TELEFAX: (703)603-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

```

```

Query Match 6.3%; Score 37.8; DB 1; Length 7218;
Best Local Similarity 1.7%; Pred. No. 0.038; Indels 0; Gaps 0;
Matches 3; Conservative 116; Mismatches 58;

```

```

Qy 409 AAGCTGGGCGCAGAGAGAAAGTATGAATGATGAGCAGCAGACCCCACTCCAGTGTG 468
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 469 GTGCTCCAGTGTGTGAGAGTACCAAGAGATGAGGAGAGAGAGAGATGGAGAGA 528
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 529 ATGAAGAGACCCCAAGCCCAAAATCATCCAGACACGCGAGATACACCGATC 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1064
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-09-056-105-31/c
; Sequence 31, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 7130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-31

```

```

Query Match 6.3%; Score 37.6; DB 4; Length 7130;
Best Local Similarity 61.0%; Pred. No. 0.044; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 39;

```

```

Qy 483 TGAGAGTACCAAGAGATGAGGAGGAGAGAGAGATGAGAGAGATGAGAGAGACCAA 542
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3553 TGAGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3494
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 543 GCCCAAAATCATCCAGACCGGAGCGGAGTACACACCG 582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3493 GGGAGAACAGTGACAAACAGGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3454
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5
US-08-913-832A-1
; Sequence 1, Application US/08913832A
; Patent No. 6329517
; GENERAL INFORMATION:
; APPLICANT: Seelig, Hans Peter
; APPLICANT: Renz, Manfred
; TITLE OF INVENTION: DERMATOMYOSITIS-SPECIFIC AUTO-ANTIGEN

```



```

; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 789:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2118
; US-09-221-0178-789

Query Match 5.7%; Score 34; DB 4; Length 2118;
Best Local Similarity 57.5%; Pred. No. 0.34;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 147 GAGCTTCAACAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCT 206
DB 1380 GAGCTCTCTATATGTTCAGAGGAAAAATCCGCTACCGCAGCTATGTACACAATACAGGGGT 1321

QY 207 GCGGTGTCAACAGCACCGAGTTCCTGGGGAGGAAATGAAGTTCCTAT 252
DB 1320 TCGTCGGCAAGTGACGGAAATCCTTCCCGAGAAAAATCACTTTTT 1275

RESULT 10
US-08-676-967-2
; Sequence 2, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,967
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2277 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-676-967-2

Query Match 5.6%; Score 33.2; DB 1; Length 2277;  
Best Local Similarity 35.9%; Pred. No. 0.65;  
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;

QY 422 GAGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481  
DB 668 GNGARGARGARAYATGARGARGARAAAGAYGAYGAYGAYGAYGAYGARG 727  
QY 482 GTGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481  
DB 728 ARGAGGNGTNTTGAAGGARGARGARGARGARGARGARGARGARGARGARG 787  
QY 542 AGCCCAAAATCATCCAGACG 563  
DB 788 ARCCNGTNCARATHCARAARMG 809

## RESULT 11

US-08-676-974-2  
Sequence 2, Application US/08676974  
Patent No. 5770422

GENERAL INFORMATION:  
APPLICANT: COLLINS, KATHLEEN  
TITLE OF INVENTION: Human Telomerase  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Science & Technology Law Group  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,974  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UCB96-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2277 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-676-974-2

Query Match 5.6%; Score 33.2; DB 1; Length 2277;  
Best Local Similarity 35.9%; Pred. No. 0.65;  
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;  
QY 422 GAGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481

DB 668 GNGARGARGARAYATGARGARGARAAAGAYGAYGAYGAYGAYGAYGARG 727  
QY 482 GTGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481  
DB 728 ARGAGGNGTNTTGAAGGARGARGARGARGARGARGARGARGARGARGARG 787  
QY 542 AGCCCAAAATCATCCAGACG 563  
DB 788 ARCCNGTNCARATHCARAARMG 809

RESULT 12  
US-09-098-487-2  
Sequence 2, Application US/09098487  
Patent No. 5917025

GENERAL INFORMATION:  
APPLICANT: COLLINS, Kathleen  
TITLE OF INVENTION: Human Telomerase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: Science & Technology Law Group  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/098,487  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UCB96-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2277 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-098-487-2

Query Match 5.6%; Score 33.2; DB 2; Length 2277;  
Best Local Similarity 35.9%; Pred. No. 0.65;  
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;

QY 422 GAGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481  
DB 668 GNGARGARGARAYATGARGARGARAAAGAYGAYGAYGAYGAYGAYGARG 727  
QY 482 GTGAGAGTGAAGTGCATGCGACGACGACCCCGCTCCAGTGTGTGTCACGCT 481  
DB 728 ARGAGGNGTNTTGAAGGARGARGARGARGARGARGARGARGARGARGARG 787  
QY 542 AGCCCAAAATCATCCAGACG 563  
DB 788 ARCCNGTNCARATHCARAARMG 809

RESULT 13  
US-09-038-3/c  
Sequence 3, Application US/09312038A  
Patent No. 6232526

GENERAL INFORMATION:  
APPLICANT: MCELROY, DAVID  
APPLICANT: KRIZ, ALAN L.  
APPLICANT: OROZCO, EMIL  
APPLICANT: GRIFFOR, MATT  
TITLE OF INVENTION: THE MAIZE A3 PROMOTER AND RICE ACTIN 2 INTRON AND  
FILE REFERENCE: DEKM-166  
CURRENT APPLICATION NUMBER: US/09/312.038A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2289  
TYPE: DNA  
ORGANISM: Zea mays/rice  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (889)..(1825)  
OTHER INFORMATION: N = A or C or G or T  
US-09-312-038-3

Query Match 5.6%; Score 33.2; DB 4; Length 2289;  
Best Local Similarity 39.3%; Pred. No. 0.66;  
Matches 107; Conservative 21; Mismatches 144; Indels 0; Gaps 0;  
QY 301 CCCGACAAAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGTGGTGAACAAGTA 360  
DB 1545 CCCGACAAAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGTGGTGAACAAGTA 360  
QY 361 GAAGATCGACCCCGCATATAATAGATCTTTATATGCTTCATCTCCAGTGGGGCCA 420  
DB 1485 KAAACMYTTCSTATTGCGWAAACMTCMSCWTTAAACARAANSCMSCRGGGCGG 1426  
QY 421 GGAGAGAAGTATGAATGCTATGCTAGCGACAGACCCCTCCAGTGTGGTCCACGTG 480  
DB 1425 GGAGGGCTGGACATGAGAGGCGATGTTGGTTACCGAATTCGATCTCGATCGCTCACC 1366  
QY 481 TGTGAGAGTACCAAGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
DB 1365 CGAAACCGGACCCCGGAGAACCTCAGCAACCGGAGGAGGAGGAGGAGGAGGAGG 1306  
QY 541 AAGCCCAAAATCATCCAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 572  
DB 1305 CAACCCACTCTAAGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1274

RESULT 14  
US-09-640-305-1/c  
Sequence 1, Application US/09640305  
Patent No. RE37447  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Yeh, Patrice  
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Rd. 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/640,305  
FILING DATE: 16-Aug-2000

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/360,673  
FILING DATE: 06-FEB-1995  
APPLICATION NUMBER: WO PCT/FR93/00623  
FILING DATE: 23-JUN-1993  
APPLICATION NUMBER: FR 92/07785  
FILING DATE: 25-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Julie K.  
REGISTRATION NUMBER: 38,619  
REFERENCE/DOCKET NUMBER: ST92040-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1685 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Kluyveromyces lactis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1683  
OTHER INFORMATION: /product= "Protease B gene"  
/gene= "K1.PRB1"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-640-305-1  
Query Match 5.5%; Score 33; DB 1; Length 1685;  
Best Local Similarity 53.5%; Pred. No. 0.64;  
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
QY 9 TAGGGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGGCAACGATGCTCTT 68  
DB 595 TAGAATCTTCTCAACAAATTCACCAATGCGCTTCTCTGATGAATCGATACTCTT 536  
QY 69 CAGCGAAGTGGAGCCAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAGGACAC 128  
DB 535 TTGTGATAGCCACATACCAAGGGAACCGGCAATATCAAAAGAGTCTTGAATAC 476  
QY 129 CACCTTCCA 137  
DB 475 CACCTTCCA 467  
RESULT 15  
US-08-360-673-1/c  
Sequence 1, Application US/08360673  
Patent No. 5679544  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Yeh, Patrice  
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Rd. 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/360,673

FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR93/00623  
FILING DATE: 23-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 92/07785  
FILING DATE: 25-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Julie K.  
REGISTRATION NUMBER: 38,619  
REFERENCE/DOCKET NUMBER: ST92040-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1685 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Kluyveromyces fragilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1683  
OTHER INFORMATION: /product= "Protease B gene"  
OTHER INFORMATION: /gene= "K1.PRB1"  
US-08-360-673-1

Query Match 5.5%; Score 33; DB 1; Length 1685;  
Best Local Similarity 53.5%; Pred. 0.64;  
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 9 TAGGACCTTACCTACATTTTAACTCCCTGATGCTTGTGTGGCAACGATGATCTT 68  
DB 595 TAGAATCTTCTTCAACCAATTCACCCATGAGGCTTCTTGATGAATGATTAATCTT 536  
OY 69 CAGCGAAGTGAAGCAGGCGCAATTTGATGCTTCTTCAACATATGACAGACAC 128  
DB 535 TTGTGAAGTAGCCACATTAACGAGAGGAAACCGCAATATCAAAAGAGCTTGAATAC 476  
OY 129 CACCTTCCA 137  
DB 475 CACCTTCCA 467

Search completed: December 14, 2002, 18:14:01  
Job time : 39.45 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:19:20 ; Search time 33.1761 Seconds  
(without alignments)  
13100.560 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggattttgggacttag.....caccgatccaccattagctga 597

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470.6	78.8	1931	5	US-09-724-676-12748 Sequence 12748, A
2	470.6	78.8	1931	5	US-09-724-676A-12748 Sequence 12748, A
3	470.6	78.8	2050	5	US-09-724-676-12756 Sequence 12756, A
4	470.6	78.8	2050	5	US-09-724-676A-12756 Sequence 12756, A
5	470.6	78.8	2297	5	US-09-724-676-12747 Sequence 12747, A
6	470.6	78.8	2297	5	US-09-724-676A-12747 Sequence 12747, A
7	459.6	77.0	2348	6	US-10-240-965-60 Sequence 60, Appl
8	408.2	68.4	1893	5	US-09-724-676-12741 Sequence 12741, A
9	408.2	68.4	1893	5	US-09-724-676A-12741 Sequence 12741, A
10	401.6	67.3	1839	5	US-09-724-676-12743 Sequence 12743, A
11	401.6	67.3	1839	5	US-09-724-676A-12743 Sequence 12743, A
12	401.6	67.3	1876	5	US-09-724-676-12745 Sequence 12745, A
13	401.6	67.3	1876	5	US-09-724-676A-12745 Sequence 12745, A
14	350.6	58.7	911	5	US-09-724-676-12749 Sequence 12749, A
15	350.6	58.7	911	5	US-09-724-676A-12749 Sequence 12749, A
16	348.2	58.3	1094	5	US-09-724-676-12750 Sequence 12750, A
17	348.2	58.3	1094	5	US-09-724-676A-12750 Sequence 12750, A
18	348.2	58.3	1213	5	US-09-724-676-12740 Sequence 12740, A
19	348.2	58.3	1213	5	US-09-724-676A-12740 Sequence 12740, A
20	285.8	47.9	1056	5	US-09-724-676-12742 Sequence 12742, A
21	285.8	47.9	1056	5	US-09-724-676A-12742 Sequence 12742, A
22	279.2	46.8	1002	5	US-09-724-676-12744 Sequence 12744, A
23	279.2	46.8	1002	5	US-09-724-676A-12744 Sequence 12744, A
24	279.2	46.8	1039	5	US-09-724-676-12746 Sequence 12746, A
25	279.2	46.8	1039	5	US-09-724-676A-12746 Sequence 12746, A
26	273.8	45.9	442	5	US-09-513-999C-1772 Sequence 1772, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-724-676-12748  
; Sequence 12748, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; NUMBER OF SEQ ID NOS: 2000-11-28  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 12748  
; LENGTH: 1931  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc\_feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc\_feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12748

Query Match 78.8%; Score 470.6; DB 5; Length 1931;  
Best Local Similarity 87.9%; Pred. No. 3e+128;  
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy	1	ATGGATTTTAGGACTTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGCGCAACGAT	60
Db	63	ATGCATTTTAGAACTTTAACTACAGTTTATAGTCCCTGATTTGTTGTGCGCAACAGT	122
Qy	61	GATGTCCTTACGCGAAGTGAGACCCAGGCGCCAAATTTGAATCCCTCTTCAGAACATATGAC	120
Db	123	GATATCTTTCAGCGAAGTGAAACCCAGGCGCCAAATTTGAGTCCCTCTTTAGGACGATGAC	182
Qy	121	AAGGACACACTTCCAGTATTTAAAGAGCTTCAAAAGTGTCCGATAAACTTCAGCAAC	180
Db	183	AAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAAGAGTCAAGATAAACTTCAGCAAC	242
Qy	181	CCCTTATCTGAGCGGCTGCGGCTGCAGAGCTGCAGACCGAGTTCCTGGGAAGGAA	240
Db	243	CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTTGGGAAGGAA	302

27	215.8	36.1	2411	5	US-09-724-676-15329	Sequence 15329, A
28	215.8	36.1	2411	5	US-09-724-676A-15329	Sequence 15329, A
29	209.6	35.1	3159	5	US-10-290-438-1	Sequence 1, Appli
30	191	32.0	1577	5	US-09-724-676-12755	Sequence 12755, A
31	191	32.0	1577	5	US-09-724-676A-12755	Sequence 12755, A
32	191	32.0	1943	5	US-09-724-676-12752	Sequence 12752, A
33	191	32.0	1943	5	US-09-724-676A-12752	Sequence 12752, A
34	186.2	31.2	615	6	US-10-290-438-8	Sequence 8, Appli
35	155	26.0	1021	6	US-10-290-438-3	Sequence 3, Appli
36	144.6	24.2	446	6	US-10-203-138A-2101	Sequence 2101, Ap
37	123	20.6	486	6	US-10-203-138A-827	Sequence 827, App
38	84.4	14.1	123	6	US-10-203-138A-5980	Sequence 5980, Ap
39	72	12.1	305	5	US-09-724-676-12751	Sequence 12751, A
40	72	12.1	305	5	US-09-724-676A-12751	Sequence 12751, A
41	71	11.9	557	5	US-09-724-676-12753	Sequence 12753, A
42	71	11.9	557	5	US-09-724-676A-12753	Sequence 12753, A
43	70.6	11.8	85	6	US-10-203-138A-7216	Sequence 7216, Ap
44	68.6	11.5	740	5	US-09-724-676-12754	Sequence 12754, A
45	68.6	11.5	740	5	US-09-724-676A-12754	Sequence 12754, A

QY 241 ATGAGTTGTAATTTGCTCAGACTTTACACATAGGAAGTTGACACTGGCTCCGCCCAAT 300  
 DB 303 ATGAAGTTATATTTTGGCTCAGACCTTTACACATAGGAAGTTGACACTGGCTCCGCCCAAT 362  
 QY 301 CCCGACAAACAGTTCTCAGATCTCCCTCCGAGCTCTCTCCCGTTGGCTGGAAACAAGTA 360  
 DB 363 CCAGACAAAGCAATTTCTGATCTCCCTCCGAGCTCTCCCGTTGGCTGGAAACAAGTA 422  
 QY 361 GAAGATGCCACCCCGCTCATAATTAATGATCTTTATATGACATCTCCAAAGCTGGGGCCA 420  
 DB 423 GAAGATGCCACCCCGCTCATAATTAATGATCTTTATATGACATCTCCAAAGCTGGGGCCA 482  
 QY 421 GGAGAGAAGTATGAATCTGATCAGACGACAGACCCCACTCCGAGTGGTGGTCCACGTG 480  
 DB 483 GGGGAAAAGTATGAATTTGACGACGACGACGACGACGACGACGACGACGACGACGACG 542  
 QY 481 TGTGAGAGTGCACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGACC 540  
 DB 543 TGTGAGAGTGCACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGACC 599  
 QY 541 AAGCCCAAAATCATCTCAGACAGGAGACCGGAGTACACACGATCCACCTTAGCTGA 597  
 DB 600 AAGCCCAAAATTTATTCAGACAGGAGACCGGAGTACACACGATCCACCTTAGCTGA 656

## RESULT 2

US-09-724-676A-12748  
 ; Sequence 12748, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12748  
 ; LENGTH: 1931  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (11)..(11)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (18)..(18)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (23)..(23)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676A-12748

Query Match 78.8%; Score 470.6; DB 5; Length 1931;  
 Best Local Similarity 87.9%; Pred. No. 3e-128; Mismatches 69; Indels 3; Gaps 1;

QY 1 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGGCAAAAGAT 60  
 DB 63 ATGATTTTGAACCTTTACTACAGTTTAGCTCCCTGATGCTGCTGGCAAAAGAT 122  
 QY 61 GATGCTTCAGCGAAGTGAAGACCAAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120  
 DB 123 GATATCTTCAGCGAAGTGAAGACCAAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 182  
 QY 121 AAGGACCAACCTTCAGATATTTTAAGAGCTTCAAAGTGTCCGGATTAATCTCAGAAC 180  
 DB 183 AAGGACATACCTTCAGATATTTTAAGAGCTTCAAAGTGTCCGGATTAATCTCAGAAC 242  
 QY 181 CCTTATCTGACGCGAGTGCAGAGGCTGCGGCTGCACAAAGCCAGTTCTGGGGAAGAA 240  
 DB 243 CCTTCTCCGACGAGTGCAGAGGCTGCGGCTGCATTAAGACTGAGTTCTGGGGAAGAA 302

QY 241 ATGAGTTGTAATTTGCTCAGACTTTACACATAGGAAGTTGACACTGGCTCCGCCCAAT 300  
 DB 303 ATGAAGTTATATTTTGGCTCAGACCTTTACACATAGGAAGTTGACACTGGCTCCGCCCAAT 362  
 QY 301 CCCGACAAACAGTTCTCAGATCTCCCTCCGAGCTCTCTCCCGTTGGCTGGAAACAAGTA 360  
 DB 363 CCAGACAAAGCAATTTCTGATCTCCCTCCGAGCTCTCCCGTTGGCTGGAAACAAGTA 422  
 QY 361 GAAGATGCCACCCCGCTCATAATTAATGATCTTTATATGACATCTCCAAAGCTGGGGCCA 420  
 DB 423 GAAGATGCCACCCCGCTCATAATTAATGATCTTTATATGACATCTCCAAAGCTGGGGCCA 482  
 QY 421 GGAGAGAAGTATGAATCTGATCAGACGACAGACCCCACTCCGAGTGGTGGTCCACGTG 480  
 DB 483 GGGGAAAAGTATGAATTTGACGACGACGACGACGACGACGACGACGACGACGACGACG 542  
 QY 481 TGTGAGAGTGCACCAAGAGATGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGACC 540  
 DB 543 TGTGAGAGTGCACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGACC 599  
 QY 541 AAGCCCAAAATCATCTCAGACAGGAGACCGGAGTACACACGATCCACCTTAGCTGA 597  
 DB 600 AAGCCCAAAATTTATTCAGACAGGAGACCGGAGTACACACGATCCACCTTAGCTGA 656

## RESULT 3

US-09-724-676-12756  
 ; Sequence 12756, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12756  
 ; LENGTH: 2050  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (7)..(7)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (47)..(47)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676-12756

Query Match 78.8%; Score 470.6; DB 5; Length 2050;  
 Best Local Similarity 87.9%; Pred. No. 3.1e-128; Mismatches 69; Indels 3; Gaps 1;

QY 1 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGGCAAAAGAT 60  
 DB 182 ATGATTTTGAACCTTTACTACAGTTTAGCTCCCTGATGCTGCTGGCAAAAGAT 241  
 QY 61 GATGCTTCAGCGAAGTGAAGACCAAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120  
 DB 242 GATATCTTCAGCGAAGTGAAGACCAAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 301  
 QY 121 AAGGACCAACCTTCAGATATTTTAAGAGCTTCAAAGTGTCCGGATTAATCTCAGAAC 180  
 DB 302 AAGGACATACCTTCAGATATTTTAAGAGCTTCAAAGTGTCCGGATTAATCTCAGAAC 361  
 QY 181 CCTTATCTGACGCGAGTGCAGAGGCTGCGGCTGCACAAAGCCAGTTCTGGGGAAGAA 240  
 DB 362 CCTTCTCCGACGAGTGCAGAGGCTGCGGCTGCATTAAGACTGAGTTCTGGGGAAGAA 421  
 QY 241 ATGAGTTGTAATTTGCTCAGACTTTACACATAGGAAGTTGACACTGGCTCCGCCCAAT 300

Db 422 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCAGACCTGGCTCCGCCAAAT 481  
Qy 301 CCCGACAAACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCCGTTGGCTGGAACAAAGTA 360  
Db 482 CCAGACAAACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCCGTTGGCTGGAACAAAGTA 541  
Qy 361 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTTATATGCACTCCAGCTGGGCGCA 420  
Db 542 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTTATATGCACTCCAGCTGGGCGCA 601  
Qy 421 GGAGAGAAGTATGAATGTCATGTCAGCGACAGACCCCACTCCAGTGTGGTCCACGTG 480  
Db 602 GGGGAAAAGTATGAATTTGACGCGAGCTGACACCACTCCAGCTGGTCCATGTA 661  
Qy 481 TGTGAGAGTACCAAGAGATCAGGAGAGGAGAGAGATGGAGAGATGAAGAGACCC 540  
Db 662 TGTGAGAGTATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718  
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597  
Db 719 AAGCCCAAAATTTATCCAGACAGGAGCGGAGTACACCGATCCACCTTAGCTGA 775

## RESULT 4

US-09-724-676A-12756  
; Sequence 12756, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12756  
; LENGTH: 2050  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc feature  
; LOCATION: (47)..(47)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12756

Query Match 78.8%; Score 470.6; DB 5; Length 2050;  
Best Local Similarity 87.9%; Pred. No. 3.1e-128;  
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy 1 ATGGATTTTAGGACCTTTAGCTACAAATTTTAGCTCCCTGATTCCTTTGTGGCAACAGAT 60  
Db 182 ATGATTTTATAGAACTTTAACTACAGTTTATGCTCCCTGATTCCTTTGTGGCAACAGAT 241  
Qy 61 GATGCTTTCAGCGAAGTACAGCGGCGCAATTTGAATCCCTTTTCAGACATATGAC 120  
Db 242 GATATCTTCAGCGAAGTACAGCGGCGCAATTTGAATCCCTTTTCAGACATATGAC 301  
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGCTGTCGCGATAAATTCAGCAAC 180  
Db 302 AAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAAAAGCTGTCGCGATAAATTCAGCAAC 361  
Qy 181 CCCTTATTCGACCGATGCCAGCTGCGGCTGCACAAAGCCGAGTTCCTGGGGAGAGAA 240  
Db 362 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 421  
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACACATAGGAAGTTTCAACCTGGCTCCGCCAAAT 300  
Db 422 ATGAAGTTATATTTTGTCTCAGACTTTTACACATAGGAAGCTTCACCTGGCTCCGCCAAAT 481  
Qy 301 CCCGACAAACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCCGTTGGCTGGAACAAAGTA 360

Db 482 CCAGACAAACAGTTTCTGATCTCCCTCCGCGCTCTCCCGTTGGCTGGAACAAAGTG 541  
Qy 361 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTTATATGCACTCCAGCTGGGCGCA 420  
Db 542 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTTATATGCACTCCAGCTGGGCGCA 601  
Qy 421 GGAGAGAAGTATGAATGTCATGTCAGCGACAGACCCCACTCCAGTGTGGTCCACGTG 480  
Db 602 GGGGAAAAGTATGAATTTGACGCGAGCTGACACCACTCCAGCTGGTCCATGTA 661  
Qy 481 TGTGAGAGTACCAAGAGATCAGGAGAGGAGAGAGATGGAGAGATGAAGAGACCC 540  
Db 662 TGTGAGAGTATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718  
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597  
Db 719 AAGCCCAAAATTTATCCAGACAGGAGCGGAGTACACCGATCCACCTTAGCTGA 775

## RESULT 5

US-09-724-676-12747  
; Sequence 12747, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12747  
; LENGTH: 2297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; NAME/KEY: misc feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12747

Query Match 78.8%; Score 470.6; DB 5; Length 2297;  
Best Local Similarity 87.9%; Pred. No. 3.2e-128;  
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy 1 ATGGATTTTAGGACCTTTAGCTACAAATTTTAGCTCCCTGATTCCTTTGTGGCAACAGAT 60  
Db 63 ATGCAATTTTAGAACTTTTAACTACAGTTTATGCTCCCTGATTCCTTTGTGGCAACAGT 122  
Qy 61 GATGCTTTCAGCGAAGTACAGCGGCGCAATTTGAATCCCTTTTCAGACATATGAC 120  
Db 123 GATATCTTCAGCGAAGTACAGCGGCGCAATTTGAATCCCTTTTCAGACATATGAC 182  
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGCTGTCGCGATAAATTCAGCAAC 180  
Db 183 AAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAAAAGCTGTCGCGATAAATTCAGCAAC 242  
Qy 181 CCCTTATTCGACCGATGCCAGCTGCGGCTGCACAAAGCCGAGTTCCTGGGAGAGAA 240  
Db 243 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 302  
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACACATAGGAAGTTTCAACCTGGCTCCGCCAAAT 300  
Db 303 ATGAAGTTATATTTTGTCTCAGACTTTTACACATAGGAAGCTTCACCTGGCTCCGCCAAAT 362

Qy	301	CCGACAAACAGTTCCTCCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAACAGTA	366
Dy	363	CCAGACAAGCAGTTCTGTATCTCCCTCCGCGCTCTCCGACAGTGGGATGGAAACAAAGTG	422
Qy	361	GAGATGCGACCCCGCGTGATAAATTCATCGATCTTTTATATGCAATCTCCAGCTGGGCGCA	420
Dy	423	GAGATGCGACCCCGCGTGATAAATTCATCTTTATATGCAATCTCCAGCTGGGCGCA	482
Qy	421	CGAGGAGATGATGAACTGCATGTCAGCGGCAACAACCCCACTCCAGCTGTGTGTCCACTG	480
Dy	483	GGGAAAAGTATGAATTGCACGACGACACTACACACACTCCACGCTGTGTGTCCATGTA	542
Qy	481	TGTGAGACTGACCAAGAGATGAGGAGAAAGAGACAGATGAGAGAGATGAAAGACCC	540
Dy	543	TGTGAGACTGATCAAGAGAGAGAGAGAGAAAGAGAA--ATGAAAGAAATGAGAGACTT	599
Qy	541	AAGCCCAAAATCATCTCAGACACGAGACCGGATGTACACCGATTCACCTTAGCTGA	597
Dy	600	AAGCCCAAAATTCATCCAGACCAAGAGAGGCGGAGATACACGCCGATTCACCTCAGCTGA	656

```

1      RESULT 6
2      US-09-724-676A-12747
3      ; Sequence 12747; Application US/09724676A
4      ; GENERAL INFORMATION:
5      ; APPLICANT: Compugen LTD
6      ; TITLE OF INVENTION: Variants of alternative splicing
7      ; PILE REFERENCE: 129181.4 Compugen
8      ; CURRENT APPLICATION NUMBER: US/09/724,676A
9      ; CURRENT FILING DATE: 2000-11-28
10     ; NUMBER OF SEQ ID NOS: 97222
11     ; SOFTWARE: PatentIn version 3.2
12     ; SEQ ID NO 12747
13     ; LENGTH: 2297
14     ; TYPE: DNA
15     ; ORGANISM: Homo sapiens
16     ; FEATURE:
17     ; NAME/KEY: misc_feature
18     ; LOCATION: (11)..(11)
19     ; OTHER INFORMATION: n is a,c,g, or t
20     ; FEATURE:
21     ; NAME/KEY: misc_feature
22     ; LOCATION: (18)..(18)
23     ; OTHER INFORMATION: n is a,c,g, or t
24     ; FEATURE:
25     ; NAME/KEY: misc_feature
26     ; LOCATION: (23)..(23)
27     ; OTHER INFORMATION: n is a,c,g, or t
28     ; US-09-724-676A-12747

```

Query Match	78.8%	Score 470.6	DB 5	Length 2257
Best Local Similarity	87.9%	Pred. No. 3.2e-128		
Matches 525; Conservative	0	Mismatches 69	Indels 3	Gaps 1

Qy	1	ATGAAATTTAGGAGCCTTACGTACAAATTTAGCTCCGTATGCTGTGGTGGCAAAAGAT	60
Db	63	ATGCATTTTGAACCTTTACTACAGTTTACTCTCCGATTTGCTGTGTGGCAAAAGT	122
Qy	61	GATGTCCTTCAGCGAAAGTAGACACGAGCCAAATTGAAATCCCTCTTCAGAACATATGAC	120
Db	123	GATATCTTCAGCGAAAGTAGAACCGAGGCCAAATTTTGAGTCCCTCTTTAGGACCTATGAC	182
Qy	121	AAGGACACACCTTCACGATATTTAAGAGCTTCMAACGTGCCGATTAACCTCAGCAAC	180
Db	183	AAGGACATACCTTCAGTATTTTAAGAGCTTCMAACGATCAGATTAACCTTAGCAAC	242
Qy	181	CCCTTATCTGACAGCCGATCCAGGCTGCGGCTGCAAGACCGAGTTCCTGGGGAAGGAA	240
Db	243	CCCTTCTCCGACAGATCCAGGCTCCACAGCTGCAATTAAGATCTTGAGTTCTTGGGAAAGAA	302
Qy	241	ATGAAATTTGATTTTGTCTACAGCTTTACACATAGGAAGTTCAACCTTGGGCTCGGCCCAAT	300
Db	303	ATGAAATTTATTTTGTCTACAGCTTTACACATAGGAAGCTACACCTTGGCTCCGCCCAAT	362

Oy	301	CCGACAAACAGTTCCTCGATCTCCCTCGGGCTCTCTCCGTTGGCTGGAAACAAGTA	360
Oy	363	CCAGCAGACAGTTTCTGATCTCCCTCCGCTCTCCGCAGTGGATGGAAACAAGTG	422
Oy	361	GAGATGCCACCCCGTCATAAATTACGATCTTTATATGCAATCTCCAGCTGGGGCCA	420
Oy	423	GAAATCGACCCCAAGTCATAAACTATGATCTCTTATATCCATCTCCAAAGCTGGGGCCA	482
Oy	421	GGAGAGAAAGTATGAACCTGCATGCGACGACACACCCCATCTCCAGTGTGGTGCTCCACGTG	480
Oy	483	GGGGAAGTATGAATTGTCACGCGACGCACTGCACACACTCCCAAGCGTGGTGCTCATGTA	542
Oy	481	TGTGAGATGCCAAGAGCAATGAGAGGAAGAGGAAAGATGGAGAGCAATGAAGAGACC	540
Oy	543	TGTGAGATGATCAAGAGAGAGAGGAAAGAAAGAGAA---ATGGAAAGAAATGAGAGAACTT	599
Oy	541	AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCATTCACCTTAGACTGA	597
Oy	600	AAGCCCAAAATTTATCCAGACACGAGAGCCGAGATGACAGCGCATTCACCTCAGACTGA	656

```

RESULT 7
US-10-240-965-60
/ Sequence 60, Application US/10240965
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.
/ APPLICANT: SHIFFMAN, Dov
/ APPLICANT: SOMOGYI, Roland
/ APPLICANT: LAMN, Richard M.
/ APPLICANT: SEILHAMER, Jeffrey J.
/ APPLICANT: PORTER, Gordon J.
/ APPLICANT: MIKITA, Thomas
/ APPLICANT: TAI, Julie
/ TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
/ FILE REFERENCE: PA-0025 PCT
/ CURRENT APPLICATION NUMBER: US/10/240.965
/ CURRENT FILING DATE: 2002-10-04
/ PRIOR APPLICATION NUMBER: 60/195,106
/ PRIOR FILING DATE: 2000-04-05
/ NUMBER OF SEQ ID NOS: 276
/ SOFTWARE: PERL Program
/ SEQ ID NO 60
/ LENGTH: 2348
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No: 042176.5
US-10-240-965-60

```

Query Match	77.0%	Score 459.6	DB 6	Length 2348
Best Local Similarity	87.8%	Pred. NO. 5.5e-125		
Matches 525	Conservative 0	Mismatches 69	Indels 4	Gaps 2

Qy	1	ATGCAATTTAGGACTTAGTAGACAATTTTAGCTCCCTGATGTCGTGGGGAACGAT	60
Db	151	ATGCATTTTAGAACTTTTAACCTACAGCTTTTAGCTCCCTGATGTCGTGGGAACGAT	210
Qy	61	GATGTCCTTCAGGGAAGTGAGACCAAGGCCAAATTGATTCCTCTTCAGAACATATGAC	120
Db	211	GATATCTTCAGGGAAGGGAACCAAGGCCAAATTTTAGTCCCTCTTTAGAGACGTATGAC	270
Qy	121	AAGGACACCACTTCAGATATTTTAGAGCTTCAAAGTGTCGGATTAACCTTCAGCAAC	180
Db	271	AAGGACATCACTTCAGATATTTTAGAGCTTCAAAGTGTCGAAATTAACCTTCAGCAAC	330
Qy	181	CCCTTATCTGACGCCGATATGCCAGGCTGGGGCTGACAAGACCCGAGTCTCAGGGGAAGAA	240
Db	331	CCCTTCTCCGACGACGATATGCCAGGCTCCAGCTGATTAAGACTGAGTTCTTGGGGAAGAA	390
Qy	241	ATGAACTGTATTTTGTCTGACATTTACACTTAGGAAGTTCAACCT-GGCTCCGCCAA	299

Db 391 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGGCTCCGCCAAA 450  
Qy 300 TCCGCAAAACAGTTCTCTATCTCCCTCCCGCCTCTCTCCCGTTGCTGGAACAAAGT 359  
Db 451 TCCGCAAAACAGTTTCTGATCTCCCTCCCGCCTCTCCCGCAGTGGATGGAACAAAGT 510  
Qy 360 AGAAGATGCCACCCCGTCATATAATAGATCTTTTATATGCCATCTCTCAAGCTGGGGCC 419  
Db 511 GGAAGATCCGACCCCGTCATATAATAGATCTTTTATATGCCATCTCTCAAGCTGGGGCC 570  
Qy 420 AGGAGAGATGATGAATGTCATGACGACAGACCCCACTCCCACTGCTGGTGGTCCAGGT 479  
Db 571 AGGGGAAAGATGATGATGTCAGCAGCAGCTGACACCACTCCCACTGCTGGTGGTCCAGT 630  
Qy 480 GTGTGAGAGTGAACCAAGAGATGAGGAGGAGAGAGAGATGAGAGATGAAGAGACC 539  
Db 631 ATGTGAGAGTGAACCAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687  
Qy 540 CAAGCCCAAAATCATCCAGACACCGGAGCCGAGTACACCCGATCCACCTTACCTGTA 597  
Db 688 TAAGCCAAAATTTATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 745

RESULT 8  
US-09-724-676-12741  
; Sequence 12741, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12741  
; LENGTH: 1893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-12741

Query Match 68.4%; Score 408.2; DB 5; Length 1893;  
Best Local Similarity 81.4%; Pred. No. 6.1e-110;  
Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

Qy 1 ATGAATTTAGGACTTTAGCTACAAATTTTATGCTCCCTGATGCTTTGTGGCAAAAGAT 60  
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTCCCGAGCCCAATTTGAGTCCCTCTTTAGGACGTATGAC 84  
Qy 61 GATGCTTTCAGCGAAAGTGAGACCGAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120  
Db 85 CGCGTGTTCGTGGACGGCTGTGCCGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144  
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAACGTGTCGGATAAACTTTCAGCAAC 180  
Db 145 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAAACGTGAGATTAACCTTCAGCAAC 204  
Qy 181 CCCTTATCTGAGCGGATGCCAGCTCGGCTGACAGACCGAGTTCCTGGGGAGGAA 240  
Db 205 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCAATAGACTGAGTTTCTGGGAAAGGAA 264  
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 300  
Db 265 ATGAAGTTATATTTTGTCTCAGACTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 324  
Qy 301 CCCGCAAAACAGTTCCTCATCTCCCTCCCGCCTCTCTCCCGTTGGCTGGAAACAAGTA 360  
Db 325 CCAGACAAGCAGTTCCTGATCTCCCTCCCGCCTCTCTCCCGAGTGGATGGAACAAGTG 384  
Qy 361 GAAGATGCTACCTTTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 420  
Db 385 GAAGATGCTACCTTTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 444  
Qy 421 CCCGCAAAACAGTTCCTCATCTCCCTCCCGCCTCTCTCCCGTTGGCTGGAAACAAGTA 480  
Db 445 GGGGAAAGATGATGATGTCAGCGAGCTGACACCTCCAGCGTGGTGGTCCATGTA 504  
Qy 481 TGTGAGATGACCAAGAGATGAGGAGGAGAGAGATGAGGAGATGAAGAGAGACC 540  
Db 505 TGTGAGATGATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
Qy 541 AAGCCCAAAATCATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 597  
Db 562 AAGCCCAAAATTTATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 618

Db 445 GGGGAAAGATGATGATGTCAGCGAGCTGACACCACTCCCGAGCTGGTGGTCCATGTA 504  
Qy 481 TGTGAGATGACCAAGAGATGAGGAGGAGAGAGATGAGGAGATGAAGAGAGACC 540  
Db 505 TGTGAGATGATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
Qy 541 AAGCCCAAAATCATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 597  
Db 562 AAGCCCAAAATTTATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 618

RESULT 9  
US-09-724-676A-12741  
; Sequence 12741, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12741  
; LENGTH: 1893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676A-12741

Query Match 68.4%; Score 408.2; DB 5; Length 1893;  
Best Local Similarity 81.4%; Pred. No. 6.1e-110;  
Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

Qy 1 ATGAATTTAGGACTTTAGCTACAAATTTTATGCTCCCTGATGCTTTGTGGCAAAAGAT 60  
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTCCCGAGCCCAATTTGAGTCCCTCTTTAGGACGTATGAC 84  
Qy 61 GATGCTTTCAGCGAAAGTGAGACCGAGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120  
Db 85 CGCGTGTTCGTGGACGGCTGTGCCGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144  
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAACGTGTCGGATAAACTTTCAGCAAC 180  
Db 145 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAAACGTGAGATTAACCTTCAGCAAC 204  
Qy 181 CCCTTATCTGAGCGGATGCCAGCTCGGCTGACAGACCGAGTTCCTGGGGAGGAA 240  
Db 205 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCAATAGACTGAGTTTCTGGGAAAGGAA 264  
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 300  
Db 265 ATGAAGTTATATTTTGTCTCAGACTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 324  
Qy 301 CCCGCAAAACAGTTCCTCATCTCCCTCCCGCCTCTCTCCCGTTGGCTGGAAACAAGTA 360  
Db 325 CCAGACAAGCAGTTCCTGATCTCCCTCCCGCCTCTCTCCCGAGTGGATGGAACAAGTG 384  
Qy 361 GAAGATGCTACCTTTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 420  
Db 385 GAAGATGCTACCTTTCAGACTTTTACATAGGAAGTTTCAACCTCGCTCCGCCAAT 444  
Qy 421 GAGAGAGATGATGATGTCAGCGAGCTGACACCTCCAGCGTGGTGGTCCATGTA 480  
Db 445 GGGGAAAGATGATGATGTCAGCGAGCTGACACCTCCAGCGTGGTGGTCCATGTA 504  
Qy 481 TGTGAGATGACCAAGAGATGAGGAGGAGAGAGATGAGGAGATGAAGAGAGACC 540  
Db 505 TGTGAGATGATCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
Qy 541 AAGCCCAAAATCATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 597  
Db 562 AAGCCCAAAATTTATCCAGACGAGGAGCCGAGTACACCCGATCCACCTTACCTGTA 618



QY 206 TGGCGCTGCACAAAGCCGAGTTCCTGGGAGGAAATGAAGTTGATTTTGTCTCAGACTT 265  
Db 213 TCCAGCTGCATAGACACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCT 272  
QY 266 TACACATAGGAAGTTACACCTCGCTCGGCCAATCCGCAAAACAGTTCTCTCATCTCCC 325  
Db 273 TACACATAGGAAGTTACACCTCGCTCGGCCAATCCGCAAAACAGTTTCTCATCTCCC 332  
QY 326 CTCGGGCTCTCTCCGTTGGCTGGGAACAAAGTAGAGATGCCACCCTCGCTCATAAAT 385  
Db 333 CTCGGGCTCTCTCCGCTGAGTGGATGGAAACAAGTGGAGATGCCACCCTCATAACT 392  
QY 386 ACGATCTTTATATGCCATCTCCAAGCTGGGCCAGGAGAGATGAATGCACTGCAG 445  
Db 393 ATGATCTCTTATATGCCATCTCCAAGCTGGGCCAGGAGAAAGTATGAATGCACTGCAG 452  
QY 446 CGACAGACCCCACTCCCAAGTGTGTGTCTCCACGTTGTGAGAGTGACCAAGAGAAATGAGG 505  
Db 453 CGACTGACACCACTCCCAAGCTGTGTGTCTCCATGTATGTGAGATGATCAAGAGAGGAGG 512  
QY 506 AGGAGAGGAGATGAGAGATGAAGAGACCAAGGCCAAATCATCCAGACACGGA 565  
Db 513 AAGAAGAGGAA--ATGGAAGAATGAGGAGACCTTAAGCCAAATAATATCCAGACCCAGGA 569  
QY 566 GACCGAGTACACACCGATCCACCTTAGCTGA 597  
Db 570 GGCGGAGTACGCCGATCCACCTCAGCTGA 601

## RESULT 13

US-09-724-676A-12745

; Sequence 12745, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 12745

; LENGTH: 1876

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-724-676A-12745

Query Match 67.3%; Score 401.6; DB 5; Length 1876;

Best Local Similarity 87.9%; Pred. No. 5.3e-108;

Matches 450; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

QY 86 GGGCCAAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTA 145  
Db 93 GGGCCAAATTTGAGTCCCTCTTTAGACCGTATGACAAAGACATCACCTTTCAGTATTTTA 152  
QY 146 AGAGCTTCAAACTGTCCGATAACTTCAGCAACCCCTATCTGAGCGGATGCCAGC 205  
Db 153 AGAGCTTCAAACTGAGTAACTTTCAGCAACCCCTTCTCCGAGCAGATGCCAGC 212  
QY 206 TGGGCTGCACAAAGACCGAGTTCTCTGGGGAAGGAATGAAGTTGTATTTTGTCTCAGACTT 265  
Db 213 TCAGCTGCATGAAGTCTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACTT 272  
QY 266 TACACATAGGAAGTTACACCTGGCTCCGCCAATCCCGCAAAACAGTTCTCATCTCCC 325  
Db 273 TACACATAGGAAGTTACACCTGGCTCCGCCAATCCCGCAAAACAGTTCTCATCTCCC 332  
QY 326 CTCGGGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCTCGCTCATAAAT 385  
Db 333 CTCGGGCTCTCTCCGCTGAGTGGATGGAAACAAGTGGAGATGCCACCCTCATAACT 392  
QY 386 ACGATCTTTATATGCCATCTCCAAGCTGGGCCAGGAGAGATGAATGCACTGCAG 445  
Db 393 ATGATCTCTTATATGCCATCTCCAAGCTGGGCCAGGAGAAAGTATGAATGCACTGCAG 452

QY 446 CGACAGACCCCACTCCCAAGTGTGTGTCTCCACGTTGTGAGAGTGACCAAGAGAAATGAGG 505  
Db 453 CGACTGACACCACTCCCAAGCTGTGTGTCTCATGTATGTGAGATGATCAAGAGAGGAGG 512  
QY 506 AGGAGAGGAGATGAGAGATGAAGAGACCAAGGCCAAATCATCCAGACACGGA 565  
Db 513 AAGAAGAGGAA--ATGGAAGAATGAGGAGACCTTAAGCCAAATAATATCCAGACCCAGGA 569  
QY 566 GACCGAGTACACACCGATCCACCTTAGCTGA 597  
Db 570 GGCGGAGTACGCCGATCCACCTCAGCTGA 601

## RESULT 14

US-09-724-676-12749

; Sequence 12749, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 12749

; LENGTH: 911

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (11)..(11)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (18)..(18)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (23)..(23)

; OTHER INFORMATION: n is a,c,g, or t

US-09-724-676-12749

Query Match 58.7%; Score 350.6; DB 5; Length 911;

Best Local Similarity 88.6%; Pred. No. 3.8e-93;

Matches 380; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGGATTTTAGGAGCTTTTAGCTTACAAATTTTAGCTCCCTGATTTGTTGTGGCAACGAT 60  
Db 63 ATGCATTTTAGAAACTTTTAACTTACAGTTTGTAGCTCCCTGATTTGCTGTGTCNAACAGT 122  
QY 61 GATGCTTTTCCGGAAGTGAGACCAAGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120  
Db 123 GATATCTTTTCCGGAAGTGAAACCAAGGCCAAATTTGAGTCCCTCTTTTAGGACGATATGAC 182  
QY 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAACTGTCGCGATATAACTTCAGCAAC 180  
Db 183 AAGGACATCACCTTTTCACTATTTTAAAGAGCTTCAAACTGAGTATTAACCTTCAGCAAC 242  
QY 181 CCTTATCTGAGCGGATGCCAGGCTCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGA 240  
Db 243 CCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGA 302  
QY 241 ATGAAGTTGATTTTGTCTCAGCTTACACATAGGAGTTTACACCTCGCTCCGCCCAAT 300  
Db 303 ATGAAGTTATATTTTGTCTCAGCTTACACATAGGAGTTTACACCTCGCTCCGCCCAAT 362  
QY 301 CCCGACAAACAGTTTCTCATCTCCCTCCGCCCTCTCCTCCGCTTGGCTGGAACAAGTA 360  
Db 363 CCAGACAGCAGTTTCTGATCTCCCTCCGCCCTCTCCGCCAGTGGGATGGAAACAAGTG 422  
QY 361 GAAAGTGCACCCCGCTCATAAATTTAGCATCTTTTATATGCCATCTCCAAGCTGGGCCA 420

Sun Dec 15 08:38:28 2002

us-09-782-953-5.rtf

**ପାଞ୍ଚ ୫**

Search completed: December 14, 2002, 19:59:19  
Job time : 36.1761 secs

Db	423	GAGAGTGGAGCCCCAGTCATAATCATATCTCTTAATATGCAATCTCCAAAGCTGGGGCCA	482
Qy	421	GGAGAGAG	429
Db	483	GCTGAGCAG	491

RESULT 15  
US-09-724-676A-12749

```

? SEQUENCE : 127497 Application US/07/072638
? GENERAL INFORMATION:
? APPLICANT : CompuGen LTD
? TITLE OF INVENTION: Variants of alternative splicing
? FILE REFERENCE: 129181.4 CompuGen
? CURRENT APPLICATION NUMBER: US/09/724,676A
? CURRENT FILING DATE: 2000-11-28
? NUMBER OF SEQ ID NOS: 97222
? SOFTWARE: Patencin version 3.2
? SEQ ID NO 13749
? LENGTH: 811

```

```

1  LENGTH: 911
2  TYPE: DNA
3  ORGANISM: Homo sapiens
4  FEATURE:
5  NAME/KEY: misc_feature
6  LOCATION: (11)-(11)
7  OTHER INFORMATION: n is a,c,g, or t
8  FEATURE:
9  NAME/KEY: misc_feature
10 LOCATION: (18)-(18)
11 OTHER INFORMATION: n is a,c,g, or t
12 FEATURE:
13 NAME/KEY: misc_feature
14 LOCATION: (23)-(23)
15 OTHER INFORMATION: n is a,c,g, or t
16
17 US-09-724-676A-12749

```

Query Match	58.7%	Score 350.6;	DB 5;	Length 911;
Best Local Similarity	88.6%;	Pred. No. 3.8e-93;		
Matches 380;	Conservative	0;	Mismatches 49;	Indels 0;
				Gaps 0.

Qy	1	ATGGATTTTAGGGACCTTAGCTACAAATTTTAGTCTCCGATGTGCTTGTGTGGCAAAACGAT	60
Db	63	ATGCATTTTATAGAAACTTAACTACAGTTTATAGCTCCCTGATTTGCTGTGTGGCAAAACAGT	122
Qy	61	GATGTCTTCAGGGAAGGAGACGAGGGCCAAATTGTAATCCCTCTTCAGAAACATATGAC	120
Db	123	GATATCTTCAGGGAAGTGAACCGAGGCCAAATTTAGTCCCTCTTATGAGAGATGAC	182
Qy	121	AAGGACACCACTTCCAGTATTTTAAAGCTTCAAACTGTGCCGATTAACCTTCAGCAAC	180
Db	183	AAGGACATCACCTTTCAGTATTTTAAAGCTTCAAAACGACCAAAATTAACCTTCAGCAAC	242
Qy	181	CCCTTATCTGCAGCCGATCCAGGCTGGCGGCTGCACAAGCCGAGTTCCTGTGGGAAGAA	240
Db	243	CCCTTCTCCGACAGCAATGCCAGGCTCAGCTGCATAAGATGATGTTCTGTGGGAAGAA	302
Qy	241	ATGAACTTATTTTGTCTCAGACTTTAACACATAGGAATTCACACCTGTGCTCCGCCAAT	300
Db	303	ATGAACTTATTTTGTCTCAGACTTTAACACATAGGAACCTCACACCTGTGCTCCGCCAAT	362
Qy	301	CCCGAACAACAGTCTCTCATCTCCCTCCGCGCTCTCTCCCGATTGTGGTGGAAACAAGTA	360
Db	363	CCAGACAAGGAAGTTTCTGATCTCCCTCCGCGCTCTCTCCCGAGTGGAGATGGAACAAGTG	422
Qy	361	GAAGATGCCACCCCCGCTCATTAATTAACGATCTTTTATATGCCATCTCCAACTGGGCCCA	420
Db	423	GAAGATGCGACCCCAAGTCATTAACATATGATCTCTTATATGCCATCTCCAACTGGGCCCA	482
Qy	421	GGAGAGAG 425	
Db	483	GGTAGCAG 491	



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 17:10:31 ; Search time 79.3082 Seconds  
(without alignments)  
11623.258 Million cell updates/sec

Title: US-09-782-953-11

Perfect score: 2331

Sequence: 1 tttttttttccccaggagt.....gggacagatgctccttac 2331

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	100.0	2331	10	US-09-782-953-11
2	2329.4	99.9	2355	12	US-10-044-090-255
3	2109	90.5	2212	10	US-09-782-953-14
4	2101	90.1	2173	10	US-09-880-107-3340
5	2094.2	89.8	2358	10	US-09-925-302-347
6	486	20.8	486	10	US-09-864-761-809
7	470.6	20.2	597	10	US-09-782-953-5
8	413.8	17.8	599	10	US-09-782-953-1
9	403.8	17.3	597	10	US-09-782-953-2
10	365.2	15.7	385	10	US-09-778-340-71
11	365.2	15.7	385	10	US-09-910-689-71
12	365.2	15.7	385	12	US-10-010-742-71
13	246.8	10.6	251	10	US-09-745-288-83
14	230.2	9.9	473	10	US-09-960-352-648
15	194.4	8.3	3184	10	US-09-954-456-497
16	194.4	8.3	3184	10	US-09-782-953-17
17	193.2	8.3	594	10	US-09-782-953-8
18	189	8.1	720	10	US-09-782-953-23
19	189	8.1	628	10	US-09-782-953-20

C 20	175	7.5	412	10	US-09-864-761-10388	Sequence 10388, A
C 21	175	7.5	446	10	US-09-864-761-2064	Sequence 2064, Ap
C 22	141.2	6.1	401	10	US-09-960-352-14684	Sequence 14684, A
C 23	139.2	6.0	380	10	US-09-960-352-1887	Sequence 1887, Ap
C 24	123	5.3	123	10	US-09-864-761-17592	Sequence 17592, A
C 25	111	4.8	111	10	US-09-864-761-27019	Sequence 27019, A
C 26	88	3.8	853	10	US-09-782-953-27	Sequence 27, Appl
C 27	85.2	3.7	272	10	US-09-960-352-11857	Sequence 11857, A
C 28	85	3.6	85	10	US-09-864-761-18808	Sequence 18808, A
C 29	81	3.5	365	10	US-09-728-445-736	Sequence 736, App
C 30	57	2.4	371	10	US-09-960-352-13037	Sequence 13037, A
C 31	51	2.2	1039	10	US-09-782-953-26	Sequence 26, Appl
C 32	41	1.8	519	10	US-09-878-574-4292	Sequence 4292, Ap
C 33	41	1.8	173808	12	US-10-003-806-10	Sequence 10, Appl
C 34	39.6	1.7	2251	10	US-09-786-256A-11	Sequence 11, Appl
C 35	39.4	1.7	419	10	US-09-960-352-11234	Sequence 1, Appl
C 36	37.6	1.6	5194	9	US-10-002-389-1	Sequence 5087, Ap
C 37	37.2	1.6	376	10	US-09-960-352-1036	Sequence 1036, Ap
C 38	37	1.6	344	10	US-09-960-352-1036	Sequence 1490, Ap
C 39	36.6	1.6	402	9	US-09-946-807-1490	Sequence 1490, Ap
C 40	36.6	1.6	402	10	US-09-795-686-1490	Sequence 1, Appl
C 41	36.6	1.6	1503841	9	US-09-946-807-1	Sequence 1, Appl
C 42	36.6	1.6	1503841	10	US-09-795-686-1	Sequence 1, Appl
C 43	36.6	1.6	1503841	10	US-09-795-686-1	Sequence 1, Appl
C 44	36.6	1.6	1503841	10	US-09-795-686-1	Sequence 1, Appl
C 45	36.4	1.6	252	10	US-09-733-523-3	Sequence 3, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-782-953-11  
; Sequence 11, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782.953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (144)..(734)  
US-09-782-953-11

Query Match 100.0%; Score 2331; DB 10; Length 2331;  
Best Local Similarity 100.0%; Pred. No 0;  
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTTTTTTTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAAGCACCAGCTCAAGA	60
Db	1	TTTTTTTTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAAGCACCAGCTCAAGA	60
Qy	61	AGGACCTACAGCCTCTGGAAAGGAATCTCACTAGGGCTTGACTGGTGGTCTGTAG	120
Db	61	AGGACCTACAGCCTCTGGAAAGGAATCTCACTAGGGCTTGACTGGTGGTCTGTAG	120
Qy	121	CGCTTTTCACTTAAGAAAGCAAGATGCACTTTTAAAGAACTTTTAACTACAGTTTACCTCCC	180
Db	121	CGCTTTTCACTTAAGAAAGCAAGATGCACTTTTAAAGAACTTTTAACTACAGTTTACCTCCC	180
Qy	181	TGATTGCTGTGTGCGAACAAGTATATCTTCAGCGAAGTGAACACCGGGCCAAATTTG	240

Db 181 TGATGGCCGTGTGGCAACAGTAGATATCTTCAGCGAAAGTGAACCGAGGCCAAATTTG 240  
Qy 241 AGTCCCTCTTTAGACGTATGACAGACATCACTTTCAGTATTTTAAGACTTCAAC 300  
Db 241 AGTCCCTCTTTAGACGTATGACAGACATCACTTTCAGTATTTTAAGACTTCAAC 300  
Qy 301 GAGTCAGAAATACTTCAAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGGTCATA 360  
Db 301 GAGTCAGAAATACTTCAAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGGTCATA 360  
Qy 361 AGACTGAGTTTCTGGGAAAGAAATGAAGTTATTTTGCTCAGACTTACATAGGAA 420  
Db 361 AGACTGAGTTTCTGGGAAAGAAATGAAGTTATTTTGCTCAGACTTACATAGGAA 420  
Qy 421 GCTCAACCTGGCTCCGCCMAATCCAGACAGCACTTCTGATCTCCCTCCGCTCTC 480  
Db 421 GCTCAACCTGGCTCCGCCMAATCCAGACAGCACTTCTGATCTCCCTCCGCTCTC 480  
Qy 481 CCCCAGTGGATGGAACAAAGTGAAGATGCCAGCCAGTCAATAATATGATCTTAT 540  
Db 481 CCCCAGTGGATGGAACAAAGTGAAGATGCCAGCCAGTCAATAATATGATCTTAT 540  
Qy 541 ATGCCATCTCCAAAGCTGSGGGCCAGGGGAAAGTATGAAATTCGACGCGACCTGA 600  
Db 541 ATGCCATCTCCAAAGCTGSGGGCCAGGGGAAAGTATGAAATTCGACGCGACCTGA 600  
Qy 601 CTCCAGCGTGTGTGCTCAATGATGTGAGAGTATGAAGAAAGGAGAAAGGAGAA 660  
Db 601 CTCCAGCGTGTGTGCTCAATGATGTGAGAGTATGAAGAAAGGAGAAAGGAGAA 660  
Qy 661 TGGAAAGAAATGAGGAGACCTTAAGCCAAATTAATCCAGACGAGGCGGAGTACAC 720  
Db 661 TGGAAAGAAATGAGGAGACCTTAAGCCAAATTAATCCAGACGAGGCGGAGTACAC 720  
Qy 721 CGATCCAGCTGAGCTGACGACGCGACGAGAGCGATTCGAAATCACTACGCGA 780  
Db 721 CGATCCAGCTGAGCTGACGACGCGACGAGAGCGATTCGAAATCACTACGCGA 780  
Qy 781 GGAATCTTTTACTGTGAGAGTGGCTGCTCAGCACTTCTCGAGAGTGGACCGGAGATG 840  
Db 781 GGAATCTTTTACTGTGAGAGTGGCTGCTCAGCACTTCTCGAGAGTGGACCGGAGATG 840  
Qy 841 GGGTGGCAGAAATCCGAGTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGCTCTT 900  
Db 841 GGGTGGCAGAAATCCGAGTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGCTCTT 900  
Qy 901 CTAAATCTGACACAGTTACTGTTCATGACCGCGGAAATGACTTGGGCAATCACTGA 960  
Db 901 CTAAATCTGACACAGTTACTGTTCATGACCGCGGAAATGACTTGGGCAATCACTGA 960  
Qy 961 GTTTGTGTGATGCGACAGGACATTTGGGACGTCTTGAGAAACAGATAAGATAGT 1020  
Db 961 GTTTGTGTGATGCGACAGGACATTTGGGACGTCTTGAGAAACAGATAAGATAGT 1020  
Qy 1021 TTTTGTACTGTTCTTTTCTGTAGGTCTGTCTGTCGAAAGGCAAGTTGATCAGTGA 1080  
Db 1021 TTTTGTACTGTTCTTTTCTGTAGGTCTGTCTGTCGAAAGGCAAGTTGATCAGTGA 1080  
Qy 1081 CTGAGAGAGAGCTTCTGTTTCTAAAGTCCCTGACAGGCGCTCTCTACTGATAGAA 1140  
Db 1081 CTGAGAGAGAGCTTCTGTTTCTAAAGTCCCTGACAGGCGCTCTCTACTGATAGAA 1140  
Qy 1141 GAGTACCAAGAAAGCGCTAGTGAGAGAGGTGTAAGAACAGACGAAATGCAATG 1200  
Db 1141 GAGTACCAAGAAAGCGCTAGTGAGAGAGGTGTAAGAACAGACGAAATGCAATG 1200  
Qy 1201 GGAATGTGAGCGTTCTCTTCTCTCTCAATGTTCTCATGTTTGTGATATATATCTG 1260  
Db 1201 GGAATGTGAGCGTTCTCTTCTCTCTCAATGTTCTCATGTTTGTGATATATATCTG 1260  
Qy 1261 ATTTACAGACTAACCTTTGTTGCTATATAAAGTTACCGGTTGTTTATCATCTTT 1320

Db 1261 ATTTACAGACTAACCTTTGTTGCTATATAAAGTTACCGGTTGTTTATCATCTTT 1320  
Qy 1321 GGGAGGCGAGAAAGGCTTTGAAACGATACCTTTCCAGATCTCGATTTCTGAC 1380  
Db 1321 GGGAGGCGAGAAAGGCTTTGAAACGATACCTTTCCAGATCTCGATTTCTGAC 1380  
Qy 1381 TCTTTGCAAGCACTTGTCTGGGAACTCTCTGCGAATGATCACTCAGATCCCA 1440  
Db 1381 TCTTTGCAAGCACTTGTCTGGGAACTCTCTGCGAATGATCACTCAGATCCCA 1440  
Qy 1441 ACCGTGCAAGTGTACTTGTGCTTTTGCAAAAGATTTATCTGAATTCCTGTAGA 1500  
Db 1441 ACCGTGCAAGTGTACTTGTGCTTTTGCAAAAGATTTATCTGAATTCCTGTAGA 1500  
Qy 1501 ATTTAGCTTATACATTCAGAGAAATAGCACTTCACTGCCAATTTAGTGGTGA 1560  
Db 1501 ATTTAGCTTATACATTCAGAGAAATAGCACTTCACTGCCAATTTAGTGGTGA 1560  
Qy 1561 TTTTATGTTAGGTGTTGGATGGACCTGACGTTCTGTTGTTCTTTATGTGTGTT 1620  
Db 1561 TTTTATGTTAGGTGTTGGATGGACCTGACGTTCTGTTGTTCTTTATGTGTGTT 1620  
Qy 1621 TCTATACATGAAATCATGCCAAATCTTTTGGAACTGTTGGTGAAGTGTGTTG 1680  
Db 1621 TCTATACATGAAATCATGCCAAATCTTTTGGAACTGTTGGTGAAGTGTGTTG 1680  
Qy 1681 TTTTACCCCAAGAGACATCAAGATACATCTGTAATTAAGCTGATGATATATGATA 1740  
Db 1681 TTTTACCCCAAGAGACATCAAGATACATCTGTAATTAAGCTGATGATATATGATA 1740  
Qy 1741 CCGTGTGATCACTTGGTGAAGATGATGCGAGTGGAGACTTAATATTAACCTACT 1800  
Db 1741 CCGTGTGATCACTTGGTGAAGATGATGCGAGTGGAGACTTAATATTAACCTACT 1800  
Qy 1801 GTGAATCATATGTTGTAAGAAAGCTGTTCCCATGTCTTAACGAGCTGAATCAAGCA 1860  
Db 1801 GTGAATCATATGTTGTAAGAAAGCTGTTCCCATGTCTTAACGAGCTGAATCAAGCA 1860  
Qy 1861 TGTCAAGTGAATAGTATGATCTGTGGCATATGAGAGGATGCAAGTGCCTTTCCCAT 1920  
Db 1861 TGTCAAGTGAATAGTATGATCTGTGGCATATGAGAGGATGCAAGTGCCTTTCCCAT 1920  
Qy 1921 TTCTGATGAGATTTGTTACTAGCTTAACATTTGTATTTTTTTTCTAAGTGTATG 1980  
Db 1921 TTCTGATGAGATTTGTTACTAGCTTAACATTTGTATTTTTTTTCTAAGTGTATG 1980  
Qy 1981 ATGTCTGTAATAGTATTAATTTTGGCTTACATACCGTAAACAATGTTGTCAATT 2040  
Db 1981 ATGTCTGTAATAGTATTAATTTTGGCTTACATACCGTAAACAATGTTGTCAATT 2040  
Qy 2041 TGAATACTTAATGCGCAAGTAAACAATGATGCTTTGAAATTTGGAAATGTTTATTC 2100  
Db 2041 TGAATACTTAATGCGCAAGTAAACAATGATGCTTTGAAATTTGGAAATGTTTATTC 2100  
Qy 2101 TTTGAGAAAGCAATATGTTGATTAATGCTTGAATGTTGATCAATCAAGAAATGAT 2160  
Db 2101 TTTGAGAAAGCAATATGTTGATTAATGCTTGAATGTTGATCAATCAAGAAATGAT 2160  
Qy 2161 AACGTTCTCAAACTGTTTACGTAAGTGTGTAAGAGGAGCGGCTTTGGAGAGCAAT 2220  
Db 2161 AACGTTCTCAAACTGTTTACGTAAGTGTGTAAGAGGAGCGGCTTTGGAGAGCAAT 2220  
Qy 2221 TGCATGCTGTCCAAAGTGTCTTGTAAAGTCTTTTAACTGAGAGGCTTAACCTCAAA 2280  
Db 2221 TGCATGCTGTCCAAAGTGTCTTGTAAAGTCTTTTAACTGAGAGGCTTAACCTCAAA 2280  
Qy 2281 ATACTTTTAACTGCAATCTTAATATAATGAGGACAGTATGCTCTTAC 2331  
Db 2281 ATACTTTTAACTGCAATCTTAATATAATGAGGACAGTATGCTCTTAC 2331

; Sequence 255, Application US/10044090  
; Patent No. US20020137081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 255  
; LENGTH: 2355  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1  
US-10-044-090-255

Query Match 99.9%; Score 2329.4; DB 12; Length 2355;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 TTTTTCCTCCAGGAGTGGGGCTGGCCCTTACTGCTTTATAGCAACAGGCTCAAGA 60  
Db 8 TTTTTCCTCCAGGAGTGGGGCTGGCCCTTACTGCTTTATAGCAACAGGCTCAAGA 67  
Qy 61 AGGAACCTACAGCTCTTGGAAAGAAATCTCACTAGGGGCTTGACTGGTGGTCTGTAG 120  
Db 68 AGGAACCTACAGCTCTTGGAAAGAAATCTCACTAGGGGCTTGACTGGTGGTCTGTAG 127  
Qy 121 CGCTTCACTCTAGAAAGCAAGATGATTTTGAACCTTTTAACTACAGTTTGTAGTCCC 180  
Db 128 CGCTTCACTCTAGAAAGCAAGATGATTTTGAACCTTTTAACTACAGTTTGTAGTCCC 187  
Qy 181 TGATTTGCTGTGTGGCAACAGATGATATCTTCAGCGAAAGTGAACCCAGGCGCAAAATTG 240  
Db 188 TGATTTGCTGTGTGGCAACAGATGATATCTTCAGCGAAAGTGAACCCAGGCGCAAAATTG 247  
Qy 241 AGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAA 300  
Db 248 AGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAA 307  
Qy 301 GAGTCAGAAATAAATCTTCAGCAACCCCTTCTCCGACGAGATCCAGGCTCCAGCTGATA 360  
Db 308 GAGTCAGAAATAAATCTTCAGCAACCCCTTCTCCGACGAGATCCAGGCTCCAGCTGATA 367  
Qy 361 AGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 420  
Db 368 AGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 427  
Qy 421 GCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 480  
Db 428 GCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 487  
Qy 481 CGCCAGTGGGATGGAACAAAGTGGAAAGATCGACCCCGAGTCATAAATGATCTCTTAT 540  
Db 488 CGCCAGTGGGATGGAACAAAGTGGAAAGATCGACCCCGAGTCATAAATGATCTCTTAT 547  
Qy 541 ATGCCATCTCAAGCTGGGCGCAGGGGAAAGATGATGAATTCACGCGAGCTGACACCA 600  
Db 548 ATGCCATCTCAAGCTGGGCGCAGGGGAAAGATGATGAATTCACGCGAGCTGACACCA 607  
Qy 601 CTCCAGCGTGGTGGTCCATGTATGTAGAGTGATCAAGAGAGGAGGAAAGAGGAA 660  
Db 608 CTCCAGCGTGGTGGTCCATGTATGTAGAGTGATCAAGAGAGGAGGAAAGAGGAA 667  
Qy 661 TGGAAAGATAGGAGACCTTAAGCAAAATTTATCCAGACAGGAGGCGGAGTACAGC 720  
Db 668 TGGAAAGATAGGAGACCTTAAGCAAAATTTATCCAGACAGGAGGCGGAGTACAGC 727  
Qy 721 CGATCCACCTCAGCTGAATCGGCAACGAGCGCATTCCTCAATCATATCTACGCGGA 780  
Db 728 CGATCCACCTCAGCTGAATCGGCAACGAGCGCATTCCTCAATCATATCTACGCGGA 787

Qy 781 GGAATCTTTTACTGTGGAGTGCTGGTCAAGACTTTCTTCGGAGGTGGCAGCGGATCG 840  
Db 788 GGAATCTTTTACTGTGGAGTGCTGGTCAAGACTTTCTTCGGAGGTGGCAGCGGATCG 847  
Qy 841 GSGTGGCAGAGAAATCCAGTTTCATGTTGCTCAGAGAGAAATCAAGGCGGTGCCCTTGT 900  
Db 848 GSGTGGCAGAGAAATCCAGTTTCATGTTGCTCAGAGAGAAATCAAGGCGGTGCCCTTGT 907  
Qy 901 CTAATGCTGCACACCACTTACTGTTTCATGGCACCCCGGGAATGACTTGGGCCAATCACTGA 960  
Db 908 CTAATGCTGCACACCACTTACTGTTTCATGGCACCCCGGGAATGACTTGGGCCAATCACTGA 967  
Qy 961 GTTTGTGTTGATCGCAACAGGACATTTGGGACTGTCTTCAGAGAAAACAGATAATGATGTG 1020  
Db 968 GTTTGTGTTGATCGCAACAGGACATTTGGGACTGTCTTCAGAGAAAACAGATAATGATGTG 1027  
Qy 1021 TTTTGTACTTTTCTGTTTCTGTTAGTGTCTGTGCGCAAGGCGAGTTGTATCAGTGAG 1080  
Db 1028 TTTTGTACTTTTCTGTTTCTGTTAGTGTCTGTGCGCAAGGCGAGTTGTATCAGTGAG 1087  
Qy 1081 CTCAGGAGAGACTTCTGTTTCTAAGTGGCTGCGAGGGGCGCACTCTCTACTGGTAGGAA 1140  
Db 1088 CTCAGGAGAGACTTCTGTTTCTAAGTGGCTGCGAGGGGCGCACTCTCTACTGGTAGGAA 1147  
Qy 1141 GAGGTACCAACAGAGAGCGCTTGTGCGAGAGAGTTGTGAAACAGAGCAATGCAATGT 1200  
Db 1148 GAGGTACCAACAGAGAGCGCTTGTGCGAGAGAGTTGTGAAACAGAGCAATGCAATGT 1207  
Qy 1201 GGAATTTGTAGCGTTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGTCATGATATACTG 1260  
Db 1208 GGAATTTGTAGCGTTTCTTCTTCTCCCTCATGTTCTCATGTTTGTGTCATGATATACTG 1267  
Qy 1261 ATTACAGACTTAACCTTTGTTTGTATATAAGTTTACACCGTTGTTTGTATACATCTTTT 1320  
Db 1268 ATTACAGACTTAACCTTTGTTTGTATATAAGTTTACACCGTTGTTTGTATACATCTTTT 1327  
Qy 1321 GGGAGCCAGAGAAAGCGTTTGGAAACGATATCACCTTCCAGATTTCTCGGATTCGAC 1380  
Db 1328 GGGAGCCAGAGAAAGCGTTTGGAAACGATATCACCTTCCAGATTTCTCGGATTCGAC 1387  
Qy 1381 TCTTTGCAACAGCACTTCTGTTCCGGAACCTTCTCTGGAATGCACTTCACTAGCATCCCA 1440  
Db 1388 TCTTTGCAACAGCACTTCTGTTCCGGAACCTTCTCTGGAATGCACTTCACTAGCATCCCA 1447  
Qy 1441 ACCGTGCAAGTGTAATCTTGTGCTTTTTCGCAAGAGTTGATCTGAAATTCCTCTGTAGA 1500  
Db 1448 ACCGTGCAAGTGTAATCTTGTGCTTTTTCGCAAGAGTTGATCTGAAATTCCTCTGTAGA 1507  
Qy 1501 ATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACTGCGCAACTTTTGTGGTGAGAAA 1560  
Db 1508 ATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACTGCGCAACTTTTGTGGTGAGAAA 1567  
Qy 1561 TTTTAGTTTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTGGTGT 1620  
Db 1568 TTTTAGTTTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTGGTGT 1627  
Qy 1621 TCTATACATGAATCATAGCCAAAACCTTTTGTGAAACTGTTTGTGGTGTGAGATGTGTTTC 1680  
Db 1628 TCTATACATGAATCATAGCCAAAACCTTTTGTGAAACTGTTTGTGGTGTGAGATGTGTTTC 1687  
Qy 1681 TTTTACCCCAAGAGACATCAAGATACACTTTGTAATTAAGCTGTATAGCATATATTCATA 1740  
Db 1688 TTTTACCCCAAGAGACATCAAGATACACTTTGTAATTAAGCTGTATAGCATATATTCATA 1747  
Qy 1741 CTTGTTGTACATTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACCT 1800  
Db 1748 CTTGTTGTACATTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACCT 1807  
Qy 1801 GTGAATCATATGTTGTAGGAAAGAGCTGTTCCCATGCTTAACAGGACTTGAATTCAGAGCA 1860  
Db 1808 GTGAATCATATGTTGTAGGAAAGAGCTGTTCCCATGCTTAACAGGACTTGAATTCAGAGCA 1867

QY 1861 TGTCAAGTGAATAGATCTGTGCGCATATGAGAGGATGACAGTCCCTTTCCCATTTCA 1920  
 Db 1868 TGTCAAGTGAATAGATCTGTGCGCATATGAGAGGATGACAGTCCCTTTCCCATTTCA 1927  
 QY 1921 TTCCTGATGGAATTTCTTACTAGGTTAACTTTTAATTTTTTTCTAGTTGTAATGTGT 1980  
 Db 1928 TTCCTGATGGAATTTCTTACTAGGTTAACTTTTAATTTTTTTCTAGTTGTAATGTGT 1987  
 QY 1981 ATGCTGGAATTAAGATTAATTTTGGCCTTAACAATCCGTAACAAATGTTTGCATTT 2040  
 Db 1988 ATGCTGGAATTAAGATTAATTTTGGCCTTAACAATCCGTAACAAATGTTTGCATTT 2047  
 QY 2041 TGAATATCTTAATGCGCACTAATGATGCTTTGGAAATTTGGAAGATGGTTTATTC 2100  
 Db 2048 TGAATATCTTAATGCGCACTAATGATGCTTTGGAAATTTGGAAGATGGTTTATTC 2107  
 QY 2101 TTTGAGAACCAATATGTTGCAATTAATGCTTTGATTTGATATCAAGAAATTTGATTG 2160  
 Db 2108 TTTGAGAACCAATATGTTGCAATTAATGCTTTGATTTGATATCAAGAAATTTGATTG 2167  
 QY 2161 AACGTTCTCAAAACCTGTTTACGGTACTGTTGTAAGAGGACCGGTTTGGAGAGACCAT 2220  
 Db 2168 AACGTTCTCAAAACCTGTTTACGGTACTGTTGTAAGAGGACCGGTTTGGAGAGACCAT 2227  
 QY 2221 TGCATGCTGTCCAAAGTGTCTGTGTAAGTGTCTTTAACTGAGAGGCTAACTCAAA 2280  
 Db 2228 TGCATGCTGTCCAAAGTGTCTGTGTAAGTGTCTTTAACTGAGAGGCTAACTCAAA 2287  
 QY 2281 ATACTTTTTTAACTGATCTATTAATAATGAGGACAGTATGCTCTTAC 2331  
 Db 2288 ATACTTTTTTAACTGATCTATTAATAATGAGGACAGTATGCTCTTAC 2338

## RESULT 3

US-09-782-953-14

Sequence 14, Application US/09782953

Patent No. US20020150953A1  
 GENERAL INFORMATION:  
 APPLICANT: WILLIAMS, R. SANDERS  
 APPLICANT: ROTHERMEL, BEVERLY  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 FILE REFERENCE: UTSP:674P21  
 CURRENT APPLICATION NUMBER: US/09/782,953  
 CURRENT FILING DATE: 2001-02-13  
 PRIOR APPLICATION NUMBER: 60/216,601  
 PRIOR FILING DATE: 2000-07-07  
 NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 14  
 LENGTH: 2212  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (25)..  
 US-09-782-953-14

Query Match 90.5%; Score 2109; DB 10; Length 2212;

Best Local Similarity 97.5%; Pred. No. 0;

Matches 2142; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 135 GAAGGCAAGATGCAATTTAGAACTTTAACTACAGTTTACCTCCCTGATTCCTGTGTG 194  
 Db 16 GACTGCGAGATGAGAGAGAGTGAAGCTGAGAGCTGCCCGGCGCAACATGGCTGTAC 75  
 QY 195 GCAAAAGTGAATATCTTACGCGAAAGTAAACAGGGGCAAAATTTGAGTCCCTCTTAAG 254  
 Db 76 CTGAGCCGGCGGTGTGTGAGAGGCTGTGCGGGCAATTTGAGTCCCTCTTAAG 135  
 QY 255 ACGTATGACAGAGACATCACTTTCAATTTTAAAGCTTCAACGAGTGAATAAC 314

Db 136 ACGTATGACAGAGACATCACTTTCAATTTTAAAGCTTCAACGAGTGAATAAC 195  
 QY 315 TTCAGCAACCCCTTCTCCGAGAGATGCAAGGCTCAGCTGATTAAGCTAGTTTCTG 374  
 Db 196 TTCAGCAACCCCTTCTCCGAGAGATGCAAGGCTCAGCTGATTAAGCTAGTTTCTG 255  
 QY 375 GGAAGGAAATGAAGTATATTTTGTCTAGACTTACATATAGAAAGCTCAACCTGGCT 434  
 Db 256 GGAAGGAAATGAAGTATATTTTGTCTAGACTTACATATAGAAAGCTCAACCTGGCT 315  
 QY 435 CCGCGAAATCCAGACAGAGCTTCTGATCTCCCTCCCGGCTCTCCGCGAGTGGATG 494  
 Db 316 CCGCGAAATCCAGACAGAGCTTCTGATCTCCCTCCCGGCTCTCCGCGAGTGGATG 375  
 QY 495 AAACAGTGAAGATGCGACCCCACTCAATTAATATATCTTTATATGCAATCTCCAG 554  
 Db 376 AAACAGTGAAGATGCGACCCCACTCAATTAATATCTTTATATGCAATCTCCAG 435  
 QY 555 CTGGGGCCAGGGGAAAGTATGAATTTGCAAGGACCACTGAACCACTCCAGCGGTG 614  
 Db 436 CTGGGGCCAGGGGAAAGTATGAATTTGCAAGGACCACTGAACCACTCCAGCGGTG 495  
 QY 615 GTCCATGATGTGAGATGATCAAGAGAGAGAGAAAGAAATGAAGAAATGAG 674  
 Db 496 GTCCATGATGTGAGATGATCAAGAGAGAGAGAAAGAAATGAAGAAATGAG 555  
 QY 675 AGACCTAAGCCAAATTAATTCAGAGCCAGAGAGCGGAGTACACGCCGATCCACCTACG 734  
 Db 556 AGACCTAAGCCAAATTAATTCAGAGCCAGAGAGCGGAGTACACGCCGATCCACCTACG 615  
 QY 735 TGAATCTGACAGCAGAGAGAGCGATTCGAATATCTCAACGGGAGAAATCTTTACTG 794  
 Db 616 TGAATCTGACAGCAGAGAGAGCGATTCGAATATCTCAACGGGAGAAATCTTTACTG 675  
 QY 795 TGAAGTGGCTGTGACAGACTTCTTCCGAGGTGCGAGCCGAGATCGGGGTGCGAAATC 854  
 Db 676 TGAAGTGGCTGTGACAGACTTCTTCCGAGGTGCGAGCCGAGATCGGGGTGCGAAATC 735  
 QY 855 CCAATTATGTTGCTCAGAGAGATCAAGGCGGTGCTCTTTCTTAAATGCTGCAAC 914  
 Db 736 CCAATTATGTTGCTCAGAGAGATCAAGGCGGTGCTCTTTCTTAAATGCTGCAAC 795  
 QY 915 CAGTTACTGTTCAATGAGCAGCCGGGAATGACTTGGGCCCAATCACTAGATGTTGTGATCG 974  
 Db 736 CAGTTACTGTTCAATGAGCAGCCGGGAATGACTTGGGCCCAATCACTAGATGTTGTGATCG 855  
 QY 975 CACAAGACATTTGGAGCTGTCTTGAGAAACAGATAATGATGTTTGTACTTGTTC 1034  
 Db 856 CACAAGACATTTGGAGCTGTCTTGAGAAACAGATAATGATGTTTGTACTTGTTC 915  
 QY 1035 TTTTCTGTAAGTCTGTCTGTGCCAAAGGCAAGTTGATCAAGTACATGAGAGAGACT 1094  
 Db 916 TTTTCTGTAAGTCTGTCTGTGCCAAAGGCAAGTTGATCAAGTACATGAGAGAGACT 975  
 QY 1095 TCCTGTTCTAAGTGGCCCTGACGGGCACTCTCTACTGTTGAGAAAGATACCAAGAGA 1154  
 Db 976 TCCTGTTCTAAGTGGCCCTGACGGGCACTCTCTACTGTTGAGAAAGATACCAAGAGA 1035  
 QY 1155 AGCCGCTAGTCAAGAGAGTGTGTAACAGACCAATGCAATGTGAAATTTGATGCT 1214  
 Db 1036 AGCCGCTAGTCAAGAGAGTGTGTAACAGACCAATGCAATGTGAAATTTGATGCT 1095  
 QY 1215 TTCCTTCTTCCCTCAATGTTTCATGTTTGGCATGATATTAAGTATTAACAAGCTAA 1274  
 Db 1096 TTCCTTCTTCCCTCAATGTTTCATGTTTGGCATGATATTAAGTATTAACAAGCTAA 1155  
 QY 1275 CTTTGTGTAATTAAGTACACGTTTGTGTTTATCATCTTTTGGAGCCAGGAAA 1334  
 Db 1156 CTTTGTGTAATTAAGTACACGTTTGTGTTTATCATCTTTTGGAGCCAGGAAA 1215  
 QY 1335 GCGTTTGAAGAAAGTATCACTTTCCAGATCTTGGAATTCCTGCACTCTTTGCAACAGA 1394  
 Db 1216 GCGTTTGAAGAAAGTATCACTTTCCAGATCTTGGAATTCCTGCACTCTTTGCAACAGA 1275

```
QY 1395 CTGCTGCGGAACCTCTTCTGGAAATGCAATTCACCTACGATCCCAACCGTGCACGGTGT 1454
Db 1276 CTGCTGCGGAACCTCTTCTGGAAATGCAATTCACCTACGATCCCAACCGTGCACGGTGT 1335
QY 1455 AACTTGTGCTTTTGGAAAAGAGTTGATCTGAAATTCCTGTGAGAAATTTAGCTTATACA 1514
Db 1336 AACTTGTGCTTTTGGAAAAGAGTTGATCTGAAATTCCTGTGAGAAATTTAGCTTATACA 1395
QY 1515 ATTACAGAAATAGCAGTTTCACTGCCAACTTTTGTAGTGGGTGAGAAATTTAGTTAGGTG 1574
Db 1396 ATTACAGAAATAGCAGTTTCACTGCCAACTTTTGTAGTGGGTGAGAAATTTAGTTAGGTG 1455
QY 1575 TTGGGATCGGACCTCAGTTTCTGTGTTCTTTTATGTGTGTTCTTATACATGAATC 1634
Db 1456 TTGGGATCGGACCTCAGTTTCTGTGTTCTTTTATGTGTGTTCTTATACATGAATC 1515
QY 1635 ATAGCCAAAACCTTTTGGAAAAGAGTTGATCTGAAATTCCTGTGAGAAATTTAGCTTATACA 1694
Db 1516 ATAGCCAAAACCTTTTGGAAAAGAGTTGATCTGAAATTCCTGTGAGAAATTTAGCTTATACA 1575
QY 1695 GACATCAAGATACACTTGTAAATGAAGCTGATACATATATTCATACCTGTGTGACACTT 1754
Db 1576 GACATCAAGATACACTTGTAAATGAAGCTGATACATATATTCATACCTGTGTGACACTT 1635
QY 1755 GGGTCAAAAGATGCGAGTGGGAGACTAAGATGATTAACCTACCTGTGAATCATATGTT 1814
Db 1636 GGGTCAAAAGATGCGAGTGGGAGACTAAGATGATTAACCTACCTGTGAATCATATGTT 1695
QY 1815 GTAGAAAAGCTGTTCCCATGCTCTAACAGGACTTGAATTCANAGCATGTCAAGTGGATG 1874
Db 1696 GTAGAAAAGCTGTTCCCATGCTCTAACAGGACTTGAATTCANAGCATGTCAAGTGGATG 1755
QY 1875 TAGATCTGGCGGATATGAGAGGATGAGTGGCTTTTCCCATTCATTCTCGATGGAAT 1934
Db 1756 TAGATCTGGCGGATATGAGAGGATGAGTGGCTTTTCCCATTCATTCTCGATGGAAT 1815
QY 1935 GTTATCTAGGTTAAACATTTGTAATTTTCTAGTTGTAATGTGATGTCGTAAATA 1994
Db 1816 GTTATCTAGGTTAAACATTTGTAATTTTCTAGTTGTAATGTGATGTCGTAAATA 1875
QY 1995 GGTATATATTTTGGCCCTTACAATACCGTAAACAAATGTTTCTAGTTGTAATGTGATGTCGTAAATA 2054
Db 1876 GGTATATATTTTGGCCCTTACAATACCGTAAACAAATGTTTCTAGTTGTAATGTGATGTCGTAAATA 1935
QY 2055 CCAAGTAAACATGATGCTTTGAAATTTGGAAGATGTTTATCTTTGAGAAAGCAAT 2114
Db 1936 CCAAGTAAACATGATGCTTTGAAATTTGGAAGATGTTTATCTTTGAGAAAGCAAT 1995
QY 2115 ATGTTTGCATTAAATGCTTTGATTTGTTTATATCAAGAAATGATGAAAGCTTCTCAACC 2174
Db 1996 ATGTTTGCATTAAATGCTTTGATTTGTTTATATCAAGAAATGATGAAAGCTTCTCAACC 2055
QY 2175 CTGTTTACGGTACTTGGTAAAGAGGAGCGGTTTGGGAGAGACCATTTGATCGCTGTTCA 2234
Db 2056 CTGTTTACGGTACTTGGTAAAGAGGAGCGGTTTGGGAGAGACCATTTGATCGCTGTTCA 2115
QY 2235 AGTGTCTTCTGTTAGTCTTTTAAACTGAGAGGCTTAACCTCAAAATATCTTTTAAAC 2294
Db 2116 AGTGTCTTCTGTTAGTCTTTTAAACTGAGAGGCTTAACCTCAAAATATCTTTTAAAC 2175
QY 2295 TGCATTCTATAATAATGGGACAGTATGCTCTTAC 2331
Db 2176 TGCATTCTATAATAATGGGACAGTATGCTCTTAC 2212
```

RESULT 4  
US-09-880-107-3340  
; Sequence 3340. Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.

```
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3340  
; LENGTH: 2173  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833  
US-09-880-107-3340
```

```
Query Match 90.1%; Score 2101; DB 10; Length 2173;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 231 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTCAGTATTTTAAAG 290  
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTCAGTATTTTAAAG 117  
QY 291 AGCTTCAAACGAGTCAAGATTAACCTTTCAGCAACCCCTCTTCGCGAGCAGATGCCAGGCTC 350  
Db 118 AGCTTCAAACGAGTCAAGATTAACCTTTCAGCAACCCCTCTTCGCGAGCAGATGCCAGGCTC 177  
QY 351 CAGCTGCATAGACGTAGTCTTTCGGAAAGGAAATGAAATTTATTTTCTCAGACCTTA 410  
Db 178 CAGCTGCATAGACGTAGTCTTTCGGAAAGGAAATGAAATTTATTTTCTCAGACCTTA 237  
QY 411 CACATAGGAAGCTCACACTGGCTCCGCCAAATCCAGACAAGCAGTTCGTATCTCCCT 470  
Db 238 CACATAGGAAGCTCACACTGGCTCCGCCAAATCCAGACAAGCAGTTCGTATCTCCCT 297  
QY 471 CCCGCTCTCCGCGAGTGGGATGGAACCAAGTGAAGATGCGACCCAGTCATAAATAT 530  
Db 298 CCCGCTCTCCGCGAGTGGGATGGAACCAAGTGAAGATGCGACCCAGTCATAAATAT 357  
QY 531 GATCTCTTATATGCCATCTCCAGCTGGGGCAGGGGAAAGTATGAATTCGACGACGC 590  
Db 358 GATCTCTTATATGCCATCTCCAGCTGGGGCAGGGGAAAGTATGAATTCGACGACGC 417  
QY 591 ACTGACACCACTCCCGAGCTGGTCCATGTATGTGAGAGTGATCAAGAGAAGGAGAA 650  
Db 418 ACTGACACCACTCCCGAGCTGGTCCATGTATGTGAGAGTGATCAAGAGAAGGAGAA 477  
QY 651 GAGAGGAAATGGAAGAAATGAGAGACCTTAAGCCAAAATTTATCCAGACCAGGAGCG 710  
Db 478 GAGAGGAAATGGAAGAAATGAGAGACCTTAAGCCAAAATTTATCCAGACCAGGAGCG 537  
QY 711 GAGTACACGCCGATCCACCTCAGCTGAACTGGGACGCGAGCAGCAGCATTCGAATCAT 770  
Db 538 GAGTACACGCCGATCCACCTCAGCTGAACTGGGACGCGAGCAGCAGCATTCGAATCAT 597  
QY 771 ACTCAGCGGAGGAAATCTTTTACTGTGGAGTGGTGTGTGCTACGACTTCTCGGAGTGGCA 830  
Db 598 ACTCAGCGGAGGAAATCTTTTACTGTGGAGTGGTGTGTGCTACGACTTCTCGGAGTGGCA 657  
QY 831 GCCGAGATCGGGTGGCAGAAATCCAGTTTCATGTTGCTCAGAGAAGAAATCAAGCGCGTG 890  
Db 658 GCCGAGATCGGGTGGCAGAAATCCAGTTTCATGTTGCTCAGAGAAGAAATCAAGCGCGTG 717  
QY 891 TCCCTTTGTTCTTAATGTGCACACCACTTACTGTTCATGCGACCCCGGATGACTTGGGC 950  
Db 718 TCCCTTTGTTCTTAATGTGCACACCACTTACTGTTCATGCGACCCCGGATGACTTGGGC 777  
QY 951 CAATCACTGAGTTTGTGGTGTGATCGCACAGGACATTTGGGACTGTCTTGAGAAAAACAGAT 1010
```

Db 778 CAATCTGAGTTTGTGATGCAAGACATTGGAGTGTGAGAAAACAGAT 837  
 QY 1011 AATGATAGTTTGTGATGTTCTTTCTGAGTTTCTGCTGCGCAAGGAGGAGTT 1070  
 Db 838 AATGATAGTTTGTGATGTTCTTTCTGAGTTTCTGCTGCGCAAGGAGGAGTT 897  
 QY 1071 GATCAGTGAAGCTCAGAGAGAGCTTCCGTTTCTTAAGGGCCGCGAGGAGGAGCTCTA 1130  
 Db 898 GATCAGTGAAGCTCAGAGAGAGCTTCCGTTTCTTAAGGGCCGCGAGGAGGAGCTCTA 957  
 QY 1131 CTGAG 1190  
 Db 958 CTGAG 1017  
 QY 1191 AATGCAATGAGAAATTTAGAGTTTCTTTCTTCCCTCATGTTCTCATGTTTGGAGT 1250  
 Db 1018 AATGCAATGAGAAATTTAGAGTTTCTTTCTTCCCTCATGTTCTCATGTTTGGAGT 1077  
 QY 1251 TATATTAAGTATTAACAAGTAACTTGTGATATTAAGTAACTTGTGATATTAACAAGT 1310  
 Db 1078 TATATTAAGTATTAACAAGTAACTTGTGATATTAAGTAACTTGTGATATTAACAAGT 1137  
 QY 1311 TACATCTTTGGAG 1370  
 Db 1138 TACATCTTTGGAG 1197  
 QY 1371 GATTCTGAGCTCTTTGCAAG 1430  
 Db 1198 GATTCTGAGCTCTTTGCAAG 1257  
 QY 1431 AGCATCCCAACAG 1490  
 Db 1258 AGCATCCCAACAG 1317  
 QY 1491 CCTCTGAGAAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1550  
 Db 1318 CCTCTGAGAAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377  
 QY 1551 GGGTGAGAAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1610  
 Db 1378 GGGTGAGAAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437  
 QY 1611 TGTGTGAGTATTAACAATTTAG 1670  
 Db 1438 TGTGTGAGTATTAACAATTTAG 1497  
 QY 1671 TAGTGTGATTTTAAACCCAG 1730  
 Db 1498 TAGTGTGATTTTAAACCCAG 1557  
 QY 1731 TATATTAACAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1790  
 Db 1558 TATATTAACAATTTAGTATTAACAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617  
 QY 1791 TAACTTACCTGAG 1850  
 Db 1618 TAACTTACCTGAG 1677  
 QY 1851 ATTCAAG 1910  
 Db 1678 ATTCAAG 1737  
 QY 1911 TCCCATTAATTTCTGAG 1970  
 Db 1738 TCCCATTAATTTCTGAG 1797  
 QY 1971 TGTATGTATGTCTGTGTAATAGATTAATTTTGGCTTAACAATACCGTAACATG 2030  
 Db 1798 TGTATGTATGTCTGTGTAATAGATTAATTTTGGCTTAACAATACCGTAACATG 1857  
 QY 2031 TTGTGATTTGAAATCTTAATGCAAGTAAACAAGATGCTTTGGAATTTGGAAGAT 2090

Db 1858 TTGTGATTTGAAATCTTAATGCAAGTAAACAATGATGCTTTGGAATTTGGAAGAT 1917  
 QY 2091 GGTATTAATTTCTGAG 2150  
 Db 1918 GGTATTAATTTCTGAG 1977  
 QY 2151 AAATGATTAAG 2210  
 Db 1978 AAATGATTAAG 2037  
 QY 2211 GAG 2270  
 Db 2038 GAG 2097  
 QY 2271 TAACTTAAATTAATTTTAACTGAGATTTAATTAATTAATTAATTAATTAATTAATTAAT 2330  
 Db 2098 TAACTTAAATTAATTTTAACTGAGATTTAATTAATTAATTAATTAATTAATTAATTAAT 2157  
 QY 2331 C 2331  
 Db 2158 C 2158

RESULT 5  
 US-09-925-302-347  
 / Sequence 347, Application US/09925302  
 / Patent No. US20020044941A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Rosen et al.  
 / TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 / FILE REFERENCE: PA104  
 / CURRENT APPLICATION NUMBER: US/09/925.302  
 / PRIOR FILING DATE: 2001-08-10  
 / PRIOR APPLICATION NUMBER: PCT/US00/05918  
 / PRIOR FILING DATE: 2000-03-08  
 / PRIOR APPLICATION NUMBER: 60/124,270  
 / NUMBER OF SEQ ID NOS: 896  
 / SOFTWARE: Patent In Ver. 2.0  
 / SEQ ID NO 347  
 / LENGTH: 2358  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-09-925-302-347

Query Match 89.8%; Score 2094.2; DB 10; Length 2358;  
 Best Local Similarity 97.3%; Pred. No. 0;  
 Matches 2137; Conservative 3; Mismatches 56; Indels 1; Gaps 1;  
 QY 135 GAAAGCAAGATGATTTTAAAGAACTTTAACTAAGTATTTAGTCCCTGATTGCGCTGTG 194  
 Db 105 GACTCGAGATGAG 164  
 QY 195 GAAAGCAAGATGATTTTAAAGAACTTTAAAGTATTTAGTCCCTGATTGCGCTGTG 254  
 Db 165 CTGAGACCCCGGCTGTGTGAG 224  
 QY 255 ACGTATGACAGAGATCACTTTGAGTATTTTAAAGAGTTTAAAGAGTCAAGTCAAGTAAAC 314  
 Db 225 ACGTATGACAGAGATCACTTTGAGTATTTTAAAGAGTTTAAAGAGTCAAGTCAAGTAAAC 284  
 QY 315 TTTCAGAACCCCTTCTCCGACAGATGCGAGGCTCCAGCTCAAGTCAAGTCAAGTCAAGT 374  
 Db 285 TTTCAGAACCCCTTCTCCGACAGATGCGAGGCTCCAGCTCAAGTCAAGTCAAGTCAAGT 344  
 QY 375 GGAAGAGAGAGAGTATTTTGTCTGAGACCTTAACATAGAGAGAGAGAGAGAGAGAGAGAG 434  
 Db 345 GGAAGAGAGAGAGTATTTTGTCTGAGACCTTAACATAGAGAGAGAGAGAGAGAGAGAGAG 404  
 QY 435 CCGCCAAATCCAGACAG 494  
 Db 405 CCG-CAAAATCCAGACAG 463





PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 809  
 LENGTH: 486  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AP000169.1  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.7  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.2  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9  
 US-09-864-761-809

Query Match 20.8%; Score 486; DB 10; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 6,2e-125;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 AAGTATGAATTGCAAGCAGCAGTGAACCACTCCCAAGCGTGTGTGCTGATGTGAG 629  
 DB 486 AAGTATGAATTGCAAGCAGCAGTGAACCACTCCCAAGCGTGTGTGCTGATGTGAG 427  
 QY 630 AGTATCAAG 689  
 DB 426 AGTATCAAG 367  
 QY 690 ATTATCCAGACAG 749  
 DB 366 ATTATCCAGACAG 307  
 QY 750 CGAGAGCCATTCCAAATCATCTCAACGGAGAGATCTTTTACTGTGAGGTGCTGCTG 809  
 DB 306 CGAGAGCCATTCCAAATCATCTCAACGGAGAGATCTTTTACTGTGAGGTGCTGCTG 247  
 QY 810 AGGACTTCTTGGAGGTGAGCCAGAGATCGGGGTGCGAGAAATCCAGTTTCTGTTGCT 869  
 DB 246 AGGACTTCTTGGAGGTGAGCCAGAGATCGGGGTGCGAGAAATCCAGTTTCTGTTGCT 187

QY 870 CAGAGAGATCAAGGCCGTGTCCTTGTCTAATGCTGCAACCACTTACTGTTTCAAG 929  
 DB 186 CAGAGAGATCAAGGCCGTGTCCTTGTCTAATGCTGCAACCACTTACTGTTTCAAG 127  
 QY 930 GCACCCGGGAATGACTTGGGCGCAATCACTGAGATTTGTGATGCGACAAGACATTTGG 989  
 DB 126 GCACCCGGGAATGACTTGGGCGCAATCACTGAGATTTGTGATGCGACAAGACATTTGG 67  
 QY 990 GACTGTCTTGAGAAAACAGATAATGATGTTTGTACTGTTCTTTTCTGTAGCTTC 1049  
 DB 66 GACTGTCTTGAGAAAACAGATAATGATGTTTGTACTGTTCTTTTCTGTAGCTTC 7  
 QY 1050 TGTCTG 1055  
 DB 6 TGTCTG 1

RESULT 7  
 US-09-782-953-5  
 Sequence 5, Application US/09782953  
 Patent No. US20020150953A1  
 GENERAL INFORMATION:  
 APPLICANT: WILLIAMS, R. SANDERS  
 APPLICANT: ROTHERMEL, BEVERLY  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 FILE REFERENCE: UTSD:674P21  
 CURRENT APPLICATION NUMBER: US/09/782,953  
 CURRENT FILING DATE: 2001-02-13  
 PRIOR APPLICATION NUMBER: 60/216,601  
 PRIOR FILING DATE: 2000-07-07  
 NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO: 5  
 LENGTH: 597  
 TYPE: DNA  
 ORGANISM: Mus musculus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)...(594)  
 US-09-782-953-5

Query Match 20.2%; Score 470.6; DB 10; Length 597;  
 Best Local Similarity 87.9%; Pred. No. 1.3e-120;  
 Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

QY 144 ATGATTTTGAAGAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACT 203  
 DB 1 ATGATTTTGAAGAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAACT 60  
 QY 204 GATATCTTCAGGAG 263  
 DB 61 GATATCTTCAGGAG 120  
 QY 264 AAGGACATCACTTTTCAATTTTAAAGCTTCAACAGAGTCAAGATTAACCTTCAGCAAC 323  
 DB 121 AAGGACATCACTTTTCAATTTTAAAGCTTCAACAGAGTCCGATTAACCTTCAGCAAC 180  
 QY 324 CCCTTCTCGCAGCAGAGATCCAGGCTCAAGTGTCAATGAAGTCAAGTGTGGAGAAAGAA 383  
 DB 181 CCCTTATCGCAGCAGAGATCCAGGCTCGAGTGTCAAGAGTCAAGTGTGGAGAAAGAA 240  
 QY 384 ATGAATTTATTTTGTCTGAGACCTTACATAGAGAGTCAACCTGCTCCGCCAAT 443  
 DB 241 ATGAATTTATTTTGTCTGAGACCTTACATAGAGAGTCAACCTGCTCCGCCAAT 300  
 QY 444 CCAGACAGAGATTTTGTGATCTCCCTCCGCTCTCCGCAAGTGGATGAGAAACAAGTG 503  
 DB 301 CCAGACAGAGATTTTGTGATCTCCCTCCGCTCTCCGCAAGTGGATGAGAAACAAGTG 360  
 QY 504 GAAGATGCGACCCCACTAATACTATGATCTTTATATGCAATCCCAAGCTGGGGGCA 563  
 DB 361 GAAGATGCGACCCCACTAATACTATGATCTTTATATGCAATCCCAAGCTGGGGGCA 420





OY 711 GAGTACAGCCGATCCACTGAGTGA 737  
 DB 571 GAGTACAGCCGATCCACTGAGTGA 597

## RESULT 10

US-09-778-320-71/C  
 ; Sequence 71, Application US/09778320  
 ; Patent No. US20010034052A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Mitcham, Jennifer  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: McNeill, Patricia D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.491C5  
 ; CURRENT APPLICATION NUMBER: US/09/778,320  
 ; CURRENT FILING DATE: 2001-02-06  
 ; NUMBER OF SEQ ID NOS: 301  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 71  
 ; LENGTH: 385  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)..(385)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-778-320-71

Query Match 15.7%; Score 365.2; DB 10; Length 385;  
 Best Local Similarity 98.2%; Pred. No. 1.7e-91;  
 Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

OY 752 AGAACCATTCCAAATCATCTCAGCGGAGGAACTTTTACTGTGAGGTGCTGTGAC 811  
 DB 385 AGGCGCATTCMAATCATCTCAGCGGAGGAACTTTTACTGTGAGGTGCTGTGAC 326  
 OY 812 GACTTCTTGGAGGTGCGACCGAGAT-CGGGGTGGCAGAAATCCAGTTTATGTTGCTC 870  
 DB 325 GACTTCTTGGAGGTGCGACCGAGATCCGGGGTGGCAGAAATCCAGTTTATGTTGCTC 266  
 OY 871 AGAAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGACACCGATTATGTTGCTC 930  
 DB 265 AGAAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGACACCGATTATGTTGCTC 206  
 OY 931 CACCGGAAATGACTTGGGCAATCATCTGAGTTGTTGATGTCACACAGACATTTGGG 990  
 DB 205 CACCGGAAATGACTTGGGCAATCATCTGAGTTGTTGATGTCACACAGACATTTGGG 146  
 OY 991 ACTGCTTGAAGAAAACAGATATAGTATGTTTGAATGTTTCTTTCTGTTGTTCT 1050  
 DB 145 ACTGCTTGAAGAAAACAGATATAGTATGTTTGAATGTTTCTTTCTGTTGTTCT 86  
 OY 1051 GTCTGTGCAAGGCGAGTTGATGAGAGCTCGAGAGAGCTTCTGTTTCTAATGTTG 1110  
 DB 85 GTCTGTGCAAGGCGAGTTGATGAGAGCTCGAGAGAGCTTCTGTTTCTAATGTTG 26  
 OY 1111 CCTGAGGGGCGACTCTACTGTT 1135  
 DB 25 CCTGAGGGGCGACTCTACTGTT 1

RESULT 11  
 US-09-910-689-71/C  
 ; Sequence 71, Application US/09910689  
 ; Patent No. US20020081609A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dillon, Davin C.

APPLICANT: Day, Craig H.  
 APPLICANT: Jiang, Yugu  
 APPLICANT: Houghton, Raymond L.  
 APPLICANT: Mitcham, Jennifer  
 APPLICANT: Wang, Tongtong  
 APPLICANT: McNeill, Patricia D.  
 APPLICANT: Harlocker, Susan L.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 FILE REFERENCE: 210121.491C6  
 CURRENT APPLICATION NUMBER: US/09/910,689  
 CURRENT FILING DATE: 2001-07-20  
 NUMBER OF SEQ ID NOS: 307  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 71  
 LENGTH: 385  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: 229, 292, 382  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-910-689-71

Query Match 15.7%; Score 365.2; DB 10; Length 385;  
 Best Local Similarity 98.2%; Pred. No. 1.7e-91;  
 Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

OY 752 AGAACCATTCCAAATCATCTCAGCGGAGGAACTTTTACTGTGAGGTGCTGTGAC 811  
 DB 385 AGGCGCATTCMAATCATCTCAGCGGAGGAACTTTTACTGTGAGGTGCTGTGAC 326  
 OY 812 GACTTCTTGGAGGTGCGACCGAGAT-CGGGGTGGCAGAAATCCAGTTTATGTTGCTC 870  
 DB 325 GACTTCTTGGAGGTGCGACCGAGATCCGGGGTGGCAGAAATCCAGTTTATGTTGCTC 266  
 OY 871 AGAAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGACACCGATTATGTTGCTC 930  
 DB 265 AGAAGAAATCAAGGCGGTGCTCCCTGTTCTAATGCTGACACCGATTATGTTGCTC 206  
 OY 931 CACCGGAAATGACTTGGGCAATCATCTGAGTTGTTGATGTCACACAGACATTTGGG 990  
 DB 205 CACCGGAAATGACTTGGGCAATCATCTGAGTTGTTGATGTCACACAGACATTTGGG 146  
 OY 991 ACTGCTTGAAGAAAACAGATATAGTATGTTTGAATGTTTCTTTCTGTTGTTCT 1050  
 DB 145 ACTGCTTGAAGAAAACAGATATAGTATGTTTGAATGTTTCTTTCTGTTGTTCT 86  
 OY 1051 GTCTGTGCAAGGCGAGTTGATGAGAGCTCGAGAGAGCTTCTGTTTCTAATGTTG 1110  
 DB 85 GTCTGTGCAAGGCGAGTTGATGAGAGCTCGAGAGAGCTTCTGTTTCTAATGTTG 26  
 OY 1111 CCTGAGGGGCGACTCTACTGTT 1135  
 DB 25 CCTGAGGGGCGACTCTACTGTT 1

RESULT 12  
 US-10-010-742-71/C  
 ; Sequence 71, Application US/10010742  
 ; Patent No. US20020146727A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Bennington, Angela Ann  
 ; APPLICANT: Zehentner, Barbara  
 ; APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.491C7  
CURRENT APPLICATION NUMBER: US/10/010.742  
CURRENT FILING DATE: 2001-11-30  
NUMBER OF SEQ ID NOS: 307  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 71  
LENGTH: 385  
TYPE: DNA  
ORGANISM: Homo sapiens  
NAME/KEY: misc feature  
LOCATION: 229, 252, 382  
OTHER INFORMATION: n = A,T,C or G  
US-10-010-742-71

Query Match 15.7%; Score 365.2; DB 12; Length 385;  
Best Local Similarity 98.2%; Pred. No. 1.7e-91;  
Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 752 AGGACGATTCCTCAATCATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTCAC 811  
DB 385 AGGCGCATTCCTCAATCATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTCAC 326  
QY 812 GACTCTTCGGAGTGGCGCCGAGAT- CGGGGTGGCAGAAATCCCAAGTTCATGTTGCTC 870  
DB 325 GACTCTTCGGAGTGGCGCCGAGAT- CGGGGTGGCAGAAATCCCAAGTTCATGTTGCTC 266  
QY 871 AGAAGAGATCAAGCCGCTGTCCTTGTCTAATGCTGCACACCACTTACTGTCATGG 930  
DB 265 AGAAGAGATCAAGCCGCTGTCCTTGTCTAATGCTGCACACCACTTACTGTCATGG 206  
QY 931 CACCCGGAATGACTTGGCCCAATCACTGAGTTTGTGGTATCGCACAAGCAATTTGGG 990  
DB 205 CACCCGGAATGACTTGGCCCAATCACTGAGTTTGTGGTATCGCACAAGCAATTTGGG 146  
QY 991 ACTGCTTGAGAAACAGATAATGATAGTGTGTTGACTTGTCTTTCTGGTAGTTCT 1050  
DB 145 ACTGCTTGAGAAACAGATAATGATAGTGTGTTGACTTGTCTTTCTGGTAGTTCT 86  
QY 1051 GTCTGCCAAGGCGAGTTGATCAGTGAGTCTCAGGAGAGAGTCTCTGTTCTTAAGTGG 1110  
DB 85 GTCTGCCAAGGCGAGTTGATCAGTGAGTCTCAGGAGAGAGTCTCTGTTCTTAAGTGG 26  
QY 1111 CTGCGAGGCGCACTCTCTACTGTT 1135  
DB 25 CTGCGAGGCGCACTCTCTACTGTT 1

RESULT 13  
US-09-745-288-83/c  
Sequence 83, Application US/09745288  
Patent No. US20010018058A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun C.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.446D1  
CURRENT APPLICATION NUMBER: US/09/745,288  
CURRENT FILING DATE: 2000-12-19  
NUMBER OF SEQ ID NOS: 101  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 83  
LENGTH: 251  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(251)

OTHER INFORMATION: n = A,T,C or G  
US-09-745-288-83

Query Match 10.6%; Score 246.8; DB 10; Length 251;  
Best Local Similarity 98.8%; Pred. No. 8.8e-59;  
Matches 248; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2081 TTGGGAAGATGTTTATTCTTTTGAGAACAATATGTTGCAATTAATGCTTTGATTGT 2140  
DB 251 TTGGGAATGTTTATTCTTTTGAGAACAATATGTTGCAATTAATGCTTTGATTGT 192  
QY 2141 TCATATCAAGAAATTCATTAACAGTTCTCAAAACCCCTGTTTACGGTACTTGGTAAGAGGGA 2200  
DB 191 TCATATCAAGAAATTCATTAACAGTTCTCAAAACCCCTGTTTACGGTACTTGGTAAGAGGGA 132  
QY 2201 GCGGTTTGGGAGAGACCAATTCATCGCTGTCCAAGTGTCTTGTAAAGTGTCTTTAAA 2260  
DB 131 GCGGTTTGGGAGAGACCAATTCATCGCTGTCCAAGTGTCTTGTAAAGTGTCTTTAAA 72  
QY 2261 CTGGAGAGCTAACCTCAAAATACTTTTTTAACTGCACTTCTATAATAATGCGCACAGT 2320  
DB 71 CTGGAGAGCTAACCTCAAAATATTTTTTAACTGCACTTCTATAATAATGCGCACAGT 12  
QY 2321 ATGCTCCTTAC 2331  
DB 11 ATGCTCCTTAC 1  
RESULT 14  
US-09-960-352-6648  
Sequence 6648, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nengbing  
APPLICANT: Byatt, John C.  
APPLICANT: Mathialagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
FILE REFERENCE: 16511.006/37-21(10298)C  
CURRENT APPLICATION NUMBER: US/09/960,352  
CURRENT FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 6648  
LENGTH: 473  
TYPE: DNA  
ORGANISM: Bos taurus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (386)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: 29-BOVMS1-009-Q1-E1-H9  
US-09-960-352-6648  
Query Match 9.9%; Score 230.2; DB 10; Length 473;  
Best Local Similarity 78.7%; Pred. No. 5.2e-54;  
Matches 354; Conservative 0; Mismatches 79; Indels 17; Gaps 6;  
QY 1426 CACTCAGCATCCCAACCGTCGCAAGTGTAACTTGTGCTTTGCAAAAGAAAGTGTATCTG 1485  
DB 24 CAATCCATCTCTCCGATGCAACCACTTGTGCTTTGCAAAAGAAAGTGTATCTG 83  
QY 1486 AAATTCCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATAGCAGTTTCTACTGCCAAT 1545  
DB 84 AAATTCCTCTGTAGCAATTCGCTTATAAAATTCAGAAATAGCAGTTTCTACTGCCAAT 143  
QY 1546 TTAGTGGGTGAGAAATTTAGTTAGGTTTGGATCGG-----ACCTCAGTTTCTGTT 1600  
DB 144 ATAGTGGGTGAGAAATTTAGTTAGTATGTTTATAGATCAGACAATACGTGGGTTTCATT 203  
QY 1601 GTTCTTTTATGTGTTTCTAT-ACATGAATCATAGCCAAAACCTTTT---TGGAA 1656  
DB 204 CTTTCTTGACGTGTAGTTTATATATACATGATCATAGCCAAACCTTTTTCGGGGGA 263



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2318.4	99.5	2348	6	US-10-240-965-60	Sequence 60, Appl	
2	2214.8	95.0	2297	5	US-09-724-676A-12747	Sequence 12747, A	
3	2214.8	95.0	2297	5	US-09-724-676A-12747	Sequence 12747, A	
4	1833.4	78.7	1943	5	US-09-724-676-12752	Sequence 12752, A	
5	1833.4	78.7	1943	5	US-09-724-676A-12752	Sequence 12752, A	
6	1575.4	67.6	2050	5	US-09-724-676-12756	Sequence 12756, A	
7	1575.4	67.6	2050	5	US-09-724-676A-12756	Sequence 12756, A	
8	1472.8	63.2	1931	5	US-09-724-676A-12748	Sequence 12748, A	
9	1472.8	63.2	1931	5	US-09-724-676-12748	Sequence 12748, A	
10	1354.4	58.1	1893	5	US-09-724-676A-12741	Sequence 12741, A	
11	1354.4	58.1	1893	5	US-09-724-676-12741	Sequence 12741, A	
12	1349.4	57.9	1876	5	US-09-724-676-12745	Sequence 12745, A	
13	1349.4	57.9	1876	5	US-09-724-676A-12745	Sequence 12745, A	
14	1348.4	57.8	1839	5	US-09-724-676-12743	Sequence 12743, A	
15	1348.4	57.8	1839	5	US-09-724-676A-12743	Sequence 12743, A	
16	1170	50.2	1577	5	US-09-724-676-12755	Sequence 12755, A	
17	1170	50.2	1577	5	US-09-724-676A-12755	Sequence 12755, A	
18	565.4	24.3	1213	5	US-09-724-676A-12740	Sequence 12740, A	
19	565.4	24.3	1213	5	US-09-724-676-12740	Sequence 12740, A	
20	486	20.8	486	6	US-10-203-138A-827	Sequence 827, App	
21	463.6	19.9	911	5	US-09-724-676-12749	Sequence 12749, A	
22	463.6	19.9	911	5	US-09-724-676A-12749	Sequence 12749, A	
23	462.8	19.9	1094	5	US-09-724-676-12750	Sequence 12750, A	
24	462.8	19.9	1094	5	US-09-724-676A-12750	Sequence 12750, A	
25	442	19.0	442	5	US-09-513-999C-1772	Sequence 1772, Ap	
26	344.4	14.8	1056	5	US-09-724-676-12742	Sequence 12742, A	

OY 241 AGTCCCTCTTTAGACGATGACAGGACATCACCTTTCAGTATTTTAAAGACTTCAAA 300  
 DB 248 AGTCCCTCTTTAGAGAGATGACAGGACATCATCTTTCAGTATTTTAAAGACTTCAAA 307  
 OY 301 GAGTCGAGATTAACCTTCAGCAACCCCTTCTCCGACAGANTCCAGGCTCCAGCTGCATA 360  
 DB 308 GAGTCGAGATTAACCTTCAGCAACCCCTTCTCCGACAGANTCCAGGCTCCAGCTGCATA 367  
 OY 361 AGACTGAGTTTCTGGGAAAGAAATGATATTTTCTGACAGCTTACACATAGGAA 420  
 DB 368 AGACTGAGTTTCTGGGAAAGAAATGATATTTTCTGACAGCTTACACATAGGAA 427  
 OY 421 GCTCAACCTT-GGCTCGGCAATCCAGACAGCAGTTTCTGATCTCCCTCCGCTCT 479  
 DB 428 GCTCAACCTTGGGCTCGGCAATCCAGACAGCAGTTTCTGATCTCCCTCCGCTCT 487  
 OY 480 CCSCCAGTGGGATGAGAAACAGATGAGATGCGACCCCACTCATAACTATGATCTCTTA 539  
 DB 488 CCGCCAGTGGGATGAGAAACAGATGAGATGCGACCCCACTCATAACTATGATCTCTTA 547  
 OY 540 TATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAAATTCGACGACGACGAC 599  
 DB 548 TATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAAATTCGACGACGACGAC 607  
 OY 600 ACTCCAGGCGTGGTGCATGATATGAGAGTGCATCAAGAGAGAGAGAGAGAGAA 659  
 DB 608 ACTCCAGGCGTGGTGCATGATATGAGAGTGCATCAAGAGAGAGAGAGAGAGAA 667  
 OY 660 ATGGAAAGATGAGAGACCTTAAGCCAAATATTCAGACCCAGAGGCGGAGTACAG 719  
 DB 668 ATGGAAAGATGAGAGACCTTAAGCCAAATATTCAGACCCAGAGGCGGAGTACAG 727  
 OY 720 CCGATCCACTCAGCTGAACTTGGCAGGCGAGAGAGCATTCCAATCATCTACAGGG 779  
 DB 728 CCGATCCACTCAGCTGAACTTGGCAGGCGAGAGAGCATTCCAATCATCTACAGGG 787  
 OY 780 AGGAATCTTTACTGTGAGAGGTGCTGTGACGACTTCTTGGAGGTGGCAGCCGAGATC 839  
 DB 788 AGGAATCTTTACTGTGAGAGGTGCTGTGACGACTTCTTGGAGGTGGCAGCCGAGATC 847  
 OY 840 GGGGTGCGAAGATCCCACTTCACTGCTCAGAGAGAAATCAAGGCGCTGTCCCTGT 889  
 DB 848 GGGGTGCGAAGATCCCACTTCACTGCTCAGAGAGAAATCAAGGCGCTGTCCCTGT 907  
 OY 900 TCTAATCTGAGACCACTTACTGTTCATGGCACCCGGAATGACTTGGGCAATCATG 959  
 DB 908 TCTAATCTGAGACCACTTACTGTTCATGGCACCCGGAATGACTTGGGCAATCATG 967  
 OY 960 AGTTGTGTGATGACACAGAGACATTTGGGACTGTCTTGAGAAACAGATATGATAGT 1019  
 DB 968 AGTTGTGTGATGACACAGAGACATTTGGGACTGTCTTGAGAAACAGATATGATAGT 1027  
 OY 1020 GTTTTGTACTGTTCTTTCTGTAGGTCTGTCTGTGCCAAGGGGCGGTTGATCAGTGA 1079  
 DB 1028 GTTTTGTACTGTTCTTTCTGTAGGTCTGTCTGTGCCAAGGGGCGGTTGATCAGTGA 1087  
 OY 1080 GCTCAGGAGAGAGCTTCTGTTCTTAAGTGGCTGACAGGGGCACTCTTACTGTTAGGA 1139  
 DB 1088 GCTCAGGAGAGAGCTTCTGTTCTTAAGTGGCTGACAGGGGCACTCTTACTGTTAGGA 1147  
 OY 1140 AGAGGTACACAGAGAGCCGCTAGTGTGAGAGAGGTTGTGAAAACAGCAGCATGCAATG 1199  
 DB 1148 AGAGGTACACAGAGAGCCGCTAGTGTGAGAGAGGTTGTGAAAACAGCAGCATGCAATG 1207  
 OY 1200 TGGAAATGTGAGCTTTCTTTCTTCTCTCATGTTCTCATGTTTGTGATATATTAAT 1259  
 DB 1208 TGGAAATGTGAGCTTTCTTTCTTCTCTCATGTTCTCATGTTTGTGATATATTAAT 1267  
 OY 1260 GATTTACAGAGATCACTTTGTTGTTATATTAAGTACACGTTGTTTACATCTTT 1319  
 DB 1268 GATTTACAGAGATCACTTTGTTGTTATATTAAGTACACGTTGTTTACATCTTT 1327  
 OY 1320 TGGGAAGCCAGGAAAGCTTTGGGAAACGATACCTTTCCAGATTTCTGGATTTCTGGA 1379

DB 1328 TGGGAAGCCAGGAAAGCTTTGGGAAACGATACCTTTCCAGATTTCTGATCTGGA 1387  
 OY 1380 CTCCTTCCACAGACATCTGCTGGGAACTCTTCTGGAATGATTCACACATCCCC 1439  
 DB 1388 CTCCTTCCACAGACATCTGCTGGGAACTCTTCTGGAATGATTCACACATCCCC 1447  
 OY 1440 AACCGTCAACGTTGAACTTGTGCTTTTGGCAAAAGAGTGTATGAAATTCCTGTAG 1499  
 DB 1448 AACCGTCAACGTTGAACTTGTGCTTTTGGCAAAAGAGTGTATGAAATTCCTGTAG 1507  
 OY 1500 AATTTAGCTTATCAATTCAGAGAAATGACAGTTTCACTGCAACTTTTATGGGTGAGAA 1559  
 DB 1508 AATTTAGCTTATCAATTCAGAGAAATGACAGTTTCACTGCAACTTTTATGGGTGAGAA 1567  
 OY 1560 AATTTAGCTTATGTTGGATGAGACCTCAGTTTCTGTGTTTCTTTATGTTGATG 1619  
 DB 1568 AATTTAGCTTATGTTGGATGAGACCTCAGTTTCTGTGTTTCTTTATGTTGATG 1627  
 OY 1620 TTTATATCATGAAATCATAGCCAAACCTTTTGGAAACTGTTGTTGAGATGATGTT 1679  
 DB 1628 TTTATATCATGAAATCATAGCCAAACCTTTTGGAAACTGTTGTTGAGATGATGTT 1687  
 OY 1680 CTTTAAACCCAGAGACATGAGATACCTTGAATTAAGCTGATAGCATATATTCAT 1729  
 DB 1688 CTTTAAACCCAGAGACATGAGATACCTTGAATTAAGCTGATAGCATATATTCAT 1747  
 OY 1740 ACTCTTATACCTTGGGTGAAAAGATAGGAGTGGAGACTAAGATATTAACCTAAC 1799  
 DB 1748 ACTCTTATACCTTGGGTGAAAAGATAGGAGTGGAGACTAAGATATTAACCTAAC 1807  
 OY 1800 TGTGAATCATATGTTGTGAGAAAGCTTCCCATGTCTTAACAGACTTGAATCAAGC 1859  
 DB 1808 TGTGAATCATATGTTGTGAGAAAGCTTCCCATGTCTTAACAGACTTGAATCAAGC 1867  
 OY 1860 AATGCAAGGATATGATGATCTGTGCGCATGAGAGAGGATGAGTCTTCCCATTC 1919  
 DB 1868 AATGCAAGGATATGATGATCTGTGCGCATGAGAGAGGATGAGTCTTCCCATTC 1927  
 OY 1920 ATTCTGATGAAATGTTATATATAGATTAACATTTGTAATTTTCTTATGTTATG 1979  
 DB 1928 ATTCTGATGAAATGTTATATATAGATTAACATTTGTAATTTTCTTATGTTATG 1987  
 OY 1980 TATGCTGTGTAATAGATTTATATTTGGCTTATACATGCTTAACATGTTTGCAT 2039  
 DB 1988 TATGCTGTGTAATAGATTTATATTTGGCTTATACATGCTTAACATGTTTGCAT 2047  
 OY 2040 TTGAATATCTTAATGCAAGTAATGATGATCTTTGGAAATTTGGAAATGATGTTTAT 2099  
 DB 2048 TTGAATATCTTAATGCAAGTAATGATGATCTTTGGAAATTTGGAAATGATGTTTAT 2107  
 OY 2100 CTTTGAAGCAATATGTTTGCATTAATGCTTTGATGTTTCTTATCAATCAAGAAATGAT 2159  
 DB 2108 CTTTGAAGCAATATGTTTGCATTAATGCTTTGATGTTTCTTATCAATCAAGAAATGAT 2167  
 OY 2160 GAACGTTTCCAAACCTGTTTACGCTGATGATGATGAGGAGACCGGTTGGAGAGACA 2219  
 DB 2168 GAACGTTTCCAAACCTGTTTACGCTGATGATGATGAGGAGACCGGTTGGAGAGACA 2227  
 OY 2220 TTGCATGCTGTCAGAGTGTCTTGTTAAGTGTCTTTTAACTGAGAGACTTAACCTCAA 2279  
 DB 2228 TTGCATGCTGTCAGAGTGTCTTGTTAAGTGTCTTTTAACTGAGAGACTTAACCTCAA 2287  
 OY 2280 AATATCTTTTAACTGATCTTAATTAATAGGCGACAGTATGCTCTTAC 2331  
 DB 2288 AATATCTTTTAACTGATCTTAATTAATAGGCGCGACAGTATGCTCTTAC 2339

RESULT 2  
 US-09-724-676-12747  
 ; Sequence 12747, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD

; TITLE UP INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 12747

; LENGTH: 2297

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (11)..(11)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (18)..(18)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (23)..(23)

; OTHER INFORMATION: n is a,c,g, or t

; US-09-724-676-12747

Query Match 95.0%; Score 2214.8; DB 5; Length 2297;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2227; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 102 TGACTCGTGGGCTCTAGCGCTTTTCACTGTAAAGAAAGCAAGATGCAATTTTAAAGAACTTT 161

Db 21 TANCCTCGTGGGCTCTAGCGCTTTTCACTGTAAAGAAAGCAAGATGCAATTTTAAAGAACTTT 80

QY 162 AACTACAGTTTGTAGTCCCTGATGCTGTGGGCAAAACAGTGATATCTTCAGCGAAAGT 221

Db 81 AACTACAGTTTGTAGTCCCTGATGCTGTGGGCAAAACAGTGATATCTTCAGCGAAAGT 140

QY 222 GAAACACAGGCGCAAAATTGAGTCCCTCTTTAGGACGTATGCAAGACATCACTTTTCAG 281

Db 141 GAAACACAGGCGCAAAATTGAGTCCCTCTTTAGGACGTATGCAAGACATCACTTTTCAG 200

QY 282 TATTTTAAAGAGCTTCAAAACAGTCAAGATAAACTTTCAGCAACCCCTTCTCCGACGAGAT 341

Db 201 TATTTTAAAGAGCTTCAAAACAGTCAAGATAAACTTTCAGCAACCCCTTCTCCGACGAGAT 260

QY 342 GCCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAAATGAAGTTATTTTGTCT 401

Db 261 GCCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAAATGAAGTTATTTTGTCT 320

QY 402 CAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGGATTTCTG 461

Db 321 CAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGGATTTCTG 380

QY 462 ATCTCCCTCCCGCTCTCCGACGTGGGATGGAACAGTGGAAAGATCGCACCCAGTC 521

Db 381 ATCTCCCTCCCGCTCTCCGACGTGGGATGGAACAGTGGAAAGATCGCACCCAGTC 440

QY 522 ATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGCGCAGGGGAAAGATGATGAATTG 581

Db 441 ATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGCGCAGGGGAAAGATGATGAATTG 500

QY 582 CACGACGAGCTGACACCACTCCAGCGTGGTGGTCCATGATGTGAGAGTCAAGAG 641

Db 501 CACGACGAGCTGACACCACTCCAGCGTGGTGGTCCATGATGTGAGAGTCAAGAG 560

QY 642 AAGGAGGAAGAGAGGAAATGGAAGATGAGGAGACCTTAAGCCAAATAATTATCCAGACC 701

Db 561 AAGGAGGAAGAGAGGAAATGGAAGATGAGGAGACCTTAAGCCAAATAATTATCCAGACC 620

QY 702 AGGAGCGGAGTACACCGGATCCACCTCAGCTGAACTGGCACCGGACGAGGACGAT 761

Db 621 AGGAGCGGAGTACACCGGATCCACCTCAGCTGAACTGGCACCGGACGAGGACGAT 680

QY 762 CCAAAATCATACTCACGGGGAGGAATCTTTTACTGTGGAGGTGGCTCAGCTTCTTCG 821

Db 681 CCAAAATCATACTCACGGGGAGGAATCTTTTACTGTGGAGGTGGCTCAGCACTTCTTCG 740

QY 822 GAGGTGGCAGCCGAGATCGGGGTGGCAGAAATCCCAAGTTCATGTTGCTCAGAGAGAATC 881

Db 741 GAGGTGGCAGCCGAGATCGGGGTGGCAGAAATCCCAAGTTCATGTTGCTCAGAGAGAATC 800

QY 882 AAGGCCGTGTCCCTTTCTAATGCTGCACACCAAGTTCATGTTTCTCATGCCACCCGGGAT 941

Db 801 AAGGCCGTGTCCCTTTCTAATGCTGCACACCAAGTTCATGTTTCTCATGCCACCCGGGAT 860

QY 942 GACTTGGGCCAATCACTGAGTTTGTGGTATCGCAACAGGACATTTGGGACATGCTCTGAG 1001

Db 861 GACTTGGGCCAATCACTGAGTTTGTGGTATCGCAACAGGACATTTGGGACATGCTCTGAG 920

QY 1002 AAAACAGATATGATGATGTTTGTACTTGTCTTTTCTGTTAGTGTCTGCTGTGCCAA 1061

Db 921 AAAACAGATATGATGATGTTTGTACTTGTCTTTTCTGTTAGTGTCTGCTGTGCCAA 980

QY 1062 GGGCAGGTTGATCAGTGAGCTCAGGAGAGGCTTCTGTTTCTTAAGTGGCTTGCAGGGGC 1121

Db 981 GGGCAGGTTGATCAGTGAGCTCAGGAGAGGCTTCTGTTTCTTAAGTGGCTTGCAGGGGC 1040

QY 1122 CACTCTCTACTGTTAGGAAGGTAACACAGGAGCCGCTAGTGCAGAGAGTTGTGAA 1181

Db 1041 CACTCTCTACTGTTAGGAAGGTAACACAGGAGCCGCTAGTGCAGAGAGTTGTGAA 1100

QY 1182 AACAGCAGCAATGCAATGCGAAATTTAGGCTTCTTCTTCTGCTCATGTTCTCATGT 1241

Db 1101 AACAGCAGCAATGCAATGCGAAATTTAGGCTTCTTCTTCTGCTCATGTTCTCATGT 1160

QY 1242 TTGTGCAATGATATTTACTGATTTTACAAGACTAACTTTTGTTCGTATATAAAGTTTACACG 1301

Db 1161 TTGTGCAATGATATTTACTGATTTTACAAGACTAACTTTTGTTCGTATATAAAGTTTACACG 1220

QY 1302 TTGTGTTTTCATCTTTTGGGAAGCCAGGAAGCGTTTGGAAACGATATCACCTTTTCCC 1361

Db 1221 TTGTGTTTTCATCTTTTGGGAAGCCAGGAAGCGTTTGGAAACGATATCACCTTTTCCC 1280

QY 1362 AGATTCTGGGATCTCGACTCTTTTGCACAGCAGCTTCTGCTGGGAACTCTTCTCTGGAATG 1421

Db 1281 AGATTCTGGGATCTCGACTCTTTTGCACAGCAGCTTCTGCTGGGAACTCTTCTCTGGAATG 1340

QY 1422 CATTCACCTCAGCATCCCCAACCGTGCACCGTAACTTGTCTTTTGCACAAAGAGTTGA 1481

Db 1341 CATTCACCTCAGCATCCCCAACCGTGCACCGTAACTTGTCTTTTGCACAAAGAGTTGA 1400

QY 1482 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCAGTTTTCACGTCCA 1541

Db 1401 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAATAGCAGTTTTCACGTCCA 1460

QY 1542 ACTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTG 1601

Db 1461 ACTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTG 1520

QY 1602 TTTCTTTTATGTGGTGTCTTATACATGAATCATAGCCAAACTTTTGTG-AAACTG 1660

Db 1521 TTTCTTTTATGTGGTGTCTTATACATGAATCATAGCCAAACTTTTGTG-AAACTG 1580

QY 1661 TTGTTTGTAGATAGTTGTTCTTTTACCCACGAGACATCAAGATACACTTGTAAATAAA 1720

Db 1581 TTGTTTGTAGATAGTTGTTCTTTTACCCACGAGACATCAAGATACACTTGTAAATAAA 1640

QY 1721 GCTGATAGCATATATTCATACCTGTTGTACCTTGGTGGAAAGTATGGCAGTGGGAGAC 1780

Db 1641 GCTGATAGCATATATTCATACCTGTTGTACCTTGGTGGAAAGTATGGCAGTGGGAGAC 1700

QY 1781 TAAAGTATTAACCTACCTGCTGAATCATATGTTGTAGGAAAGTCTTCCCATGTCTAA 1840

Db 1701 TAAAGTATTAACCTACCTGCTGAATCATATGTTGTAGGAAAGTCTTCCCATGTCTAA 1760

QY 1841 CAGGACTTGAATTCAAAGCATGTCAAAGTGAATAGATCTGTGGCGATATGAGAGGAT 1900



Db 1761 CAGACTTGATTCAAAGCATGTCAGATAGATCTGTGGATGAGAGGAT 1820  
Qy 1901 GAGAGGCTTTCCCATTCATTCCTGATGGAATGTTACTAGGTTAACTTTGAAAT 1960  
Db 1821 GAGAGGCTTTCCCATTCATTCCTGATGGAATGTTACTAGGTTAACTTTGAAAT 1880  
Qy 1961 TTTTCTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020  
Db 1881 TTTTCTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1940  
Qy 2021 CGTAAACATGTTGTCATTTTGAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAAT 2080  
Db 1941 CGTAAACATGTTGTCATTTTGAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAAT 2000  
Qy 2081 TTTGGAAGATGTTTATTTCTTTGAGAGCAATTTGATTTGATTTGATTTGATTTGATTTGAT 2140  
Db 2001 TTTGGAAGATGTTTATTTCTTTGAGAGCAATTTGATTTGATTTGATTTGATTTGATTTGAT 2060  
Qy 2141 TCATATCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2200  
Db 2061 TCATATCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2120  
Qy 2201 GCGGTTTGGAGAGACCATTCGATGCTGCTCAAGTGTTCCTTTGATGCTTTTAA 2260  
Db 2121 GCGGTTTGGAGAGACCATTCGATGCTGCTCAAGTGTTCCTTTGATGCTTTTAA 2180  
Qy 2261 CTGGAGAGGCTACCTCAAAATCTTTTAACTGATCTTAATTAATGAGGACAGT 2320  
Db 2181 CTGGAGAGGCTACCTCAAAATCTTTTAACTGATCTTAATTAATGAGGACAGT 2240  
Qy 2321 ATGCTCCTTAC 2331  
Db 2241 ATGCTCCTTAC 2251

RESULT 3  
US-09-724-676A-12747  
Sequence 12747, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variance of alternative splicing  
FILE REFERENCE: 129181, 4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724, 676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12747  
LENGTH: 2297  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (11)-(11)  
OTHER INFORMATION: n is a,c,g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (18)-(18)  
OTHER INFORMATION: n is a,c,g, or t  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (23)-(23)  
OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12747

Query Match 95.0%; Score 2214.8; DB 5; Length 2297;  
Best Local Similarity 99.8%; Pred. NO. 0;  
Matches 2227; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 102 TGACATGCGTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGATTTTAAACCTTT 161  
Db 21 TATGCTGCTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGATTTTAAACCTTT 80  
Qy 162 AACTAGAGTTTATGCTCCCTGATGCTGCTGTGTGGCAACAGTATATCTTACAGGAAAGT 221

Db 81 AACTAGAGTTTATGCTCCCTGATGCTGCTGTGTGGCAACAGTATATCTTACAGGAAAGT 140  
Qy 222 GAAACAGAGGCAAAATTTGAGTCCCTCTTTAGACGTATGACAAAGACATCACCTTTAG 281  
Db 141 GAAACAGAGGCAAAATTTGAGTCCCTCTTTAGACGTATGACAAAGACATCACCTTTAG 200  
Qy 282 TATTTTAAAGCTTCAACAGATCAGAAATTAATTTAGCAACCCCTTCTCGACAGCAT 341  
Db 201 TATTTTAAAGCTTCAACAGATCAGAAATTAATTTAGCAACCCCTTCTCGACAGCAT 260  
Qy 342 GCCAGGCTCCAGCTGCAATTAAGATGATGCTGCGAAAGAAATGAAATTAATTTAGT 401  
Db 261 GCCAGGCTCCAGCTGCAATTAAGATGATGCTGCGAAAGAAATGAAATTAATTTAGT 320  
Qy 402 CAGACTTACACATAGAAAGCTCACACCTGCTCCGCCAATTCAGACAAAGCATTTCTG 461  
Db 321 CAGACTTACACATAGAAAGCTCACACCTGCTCCGCCAATTCAGACAAAGCATTTCTG 380  
Qy 462 ATCTCCCTCCGCTCTCTCCGACAGTGGATGAAACAAAGTGAAGATGCAACCCCATG 521  
Db 381 ATCTCCCTCCGCTCTCTCCGACAGTGGATGAAACAAAGTGAAGATGCAACCCCATG 440  
Qy 522 ATAACTATGATCTCTTATATGCAATCCCAAGCTGGGCGCAGGAGAAATGATGAAATTTG 581  
Db 441 ATAACTATGATCTCTTATATGCAATCTCCAAAGCTGGGCGCAGGAGAAATGATGAAATTTG 500  
Qy 582 CACGAGAGGACTGACACCATCTCCAGCTGTGTGCTCATGATGATGATGATGATGATGAT 641  
Db 501 CACGAGAGGACTGACACCATCTCCAGCTGTGTGCTCATGATGATGATGATGATGATGAT 560  
Qy 642 AAGAGAGAAAGAGAGAAATGAAAGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAG 701  
Db 561 AAGAGAGAAAGAGAGAAATGAAAGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAG 620  
Qy 702 AAGAGAGGAGAGTACAGCGGCAATCTTAACTGTGAGGAGGCTGTGCTGCACTTTCTTG 761  
Db 621 AAGAGAGGAGAGTACAGCGGCAATCTTAACTGTGAGGAGGCTGTGCTGCACTTTCTTG 680  
Qy 762 CCMAATCATACTACAGGAGAAATCTTTTAACTGTGAGGAGGCTGTGCTGCACTTTCTTG 821  
Db 681 CCMAATCATACTACAGGAGAAATCTTTTAACTGTGAGGAGGCTGTGCTGCACTTTCTTG 740  
Qy 822 GAGGTGGAGCGGCAATCGGAGTGGGAGAAATCCAGTTCATGTTGCTCAGAGAGAAATC 881  
Db 741 GAGGTGGAGCGGCAATCGGAGTGGGAGAAATCCAGTTCATGTTGCTCAGAGAGAAATC 800  
Qy 882 AAGGCGGTGCTGCTGCTTAACTGTGCTGCAACAGATTACTGTTGATGAGCAACCGGAAAT 941  
Db 801 AAGGCGGTGCTGCTGCTTAACTGTGCTGCAACAGATTACTGTTGATGAGCAACCGGAAAT 860  
Qy 942 GACTTGGGCAATCACTGAGTTGTGTGATGTCGCAAGAGCAATTTGGAGCTGTCTTGAG 1001  
Db 861 GACTTGGGCAATCACTGAGTTGTGTGATGTCGCAAGAGCAATTTGGAGCTGTCTTGAG 920  
Qy 1002 AAAACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1061  
Db 921 AAAACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980  
Qy 1062 GGGCAGGTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTTTCTAAAGTGCCTGACAGGGGC 1121  
Db 981 GGGCAGGTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTTTCTAAAGTGCCTGACAGGGGC 1040  
Qy 1122 CACTCTCTAGTGTAGAAAGAGTACACAGAGAGCGGCTGTGTGAGAGAGTGTGA 1181  
Db 1041 CACTCTCTAGTGTAGAAAGAGTACACAGAGAGCGGCTGTGTGAGAGAGTGTGA 1100  
Qy 1182 AAGAGAGCAAGTGAAGTGAAGTGTAGCGTTCTTCTTCCCTCATGTTCTCATGT 1241  
Db 1101 AAGAGAGCAAGTGAAGTGAAGTGTAGCGTTCTTCTTCCCTCATGTTCTCATGT 1160  
Qy 1242 TTGTGATGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301



Db 1161 TTGTGATGATATTAATGATTTACAGACTAACCTTTGTTGATATATAAGTTACACG 1220  
QY 1302 TTGTTGTTTACATCTTTTGGGAAGCAGGAAAGCGTTTGGAAAGCGTATACCTTTCC 1361  
Db 1221 TTGTTGTTTACATCTTTTGGGAAGCAGGAAAGCGTTTGGAAAGCGTATACCTTTCC 1280  
QY 1362 AGATTCTCGGATCTCGACTCTTTGCAACAGCAGCTTGTCTGGGAACCTCTTCTGGAATG 1421  
Db 1281 AGATTCTCGGATCTCGACTCTTTGCAACAGCAGCTTGTCTGGGAACCTCTTCTGGAATG 1340  
QY 1422 CATTTCACTCAGCATCCCCAACCGTGCACGCTGTAACTTTGTCTGGGAACCTCTTCTGGAATG 1481  
Db 1341 CATTTCACTCAGCATCCCCAACCGTGCACGCTGTAACTTTGTCTGGGAACCTCTTCTGGAATG 1400  
QY 1482 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAATTCAGAGAATAGCAGTTTCACTGCCA 1541  
Db 1401 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAATTCAGAGAATAGCAGTTTCACTGCCA 1460  
QY 1542 ACTTTTAGTGGGTGAGAAATTTAGCTTTAGGTGTGGGATCGGACCTCAGTTTCTGTTG 1601  
Db 1461 ACTTTTAGTGGGTGAGAAATTTAGCTTTAGGTGTGGGATCGGACCTCAGTTTCTGTTG 1520  
QY 1602 TTCTCTTTATGTTGGTGTCTATACATGAATCATAGCCAAATACTTTTTGG-AAAAGT 1660  
Db 1521 TTCTCTTTATGTTGGTGTCTATACATGAATCATAGCCAAATACTTTTTGGAAAGT 1580  
QY 1661 TTGTTGATAGTTGGTGTCTTTTACCACGAAAGACATCAAGTATGAGTGTGGAATGAGAC 1720  
Db 1581 TTGTTGATAGTTGGTGTCTTTTACCACGAAAGACATCAAGTATGAGTGTGGAATGAGAC 1640  
QY 1721 GCTGATAGCATATTTATCATCTGTTGTACACTTTGGGTGAAAGTATGGCAGTGGGAGAC 1780  
Db 1641 GCTGATAGCATATTTATCATCTGTTGTACACTTTGGGTGAAAGTATGGCAGTGGGAGAC 1700  
QY 1781 TAAGATGATTTAACTACCTGTTGAATCATATGTTGTAGGAAAGCTGTTCCCATGCTAA 1840  
Db 1701 TAAGATGATTTAACTACCTGTTGAATCATATGTTGTAGGAAAGCTGTTCCCATGCTAA 1760  
QY 1841 CAGCACTTGAATTCAGAGCATGTCAGTGTAGTATGATCTGTGGCGATATGAGAGGAT 1900  
Db 1761 CAGCACTTGAATTCAGAGCATGTCAGTGTAGTATGATCTGTGGCGATATGAGAGGAT 1820  
QY 1901 GCAGTGCTTTCCCATTCATCTCTGATGGAATGTTTATATAGTTTAACTAGTTTAACTTGAAT 1960  
Db 1821 GCAGTGCTTTCCCATTCATCTCTGATGGAATGTTTATATAGTTTAACTAGTTTAACTTGAAT 1880  
QY 1961 TTTTCTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACT 2020  
Db 1881 TTTTCTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACTAGTTTAACT 1940  
QY 2021 CGTAAACAATGTTTGTCAATTTTGAATACTTAATGCCAAGTTAACTAGTTTAACTAGTTTAACT 2080  
Db 1941 CGTAAACAATGTTTGTCAATTTTGAATACTTAATGCCAAGTTAACTAGTTTAACTAGTTTAACT 2000  
QY 2081 TTTGGAAGATGTTTATCTTTTGAAGCAAAATATGTTTGAATTAATGTTTGAATTTGATTTCT 2140  
Db 2001 TTTGGAAGATGTTTATCTTTTGAAGCAAAATATGTTTGAATTAATGTTTGAATTTGATTTCT 2060  
QY 2141 TCATATCAAGAAATTCATTTGAAGCTTCTAAACCTGTTTACCGTACTTGGTAAAGGGA 2200  
Db 2061 TCATATCAAGAAATTCATTTGAAGCTTCTAAACCTGTTTACCGTACTTGGTAAAGGGA 2120  
QY 2201 GCCGTTTGGGAGAGACCATTCATCGCTGTCCAGTGTCTTCTGTTAGTGTCTTTTAA 2260  
Db 2121 GCCGTTTGGGAGAGACCATTCATCGCTGTCCAGTGTCTTCTGTTAGTGTCTTTTAA 2180  
QY 2261 CTGGAGAGCTTAACCTCAAAATCTTTTTTAACTGCAATTTCTATATAAATGGGACAGT 2320  
Db 2181 CTGGAGAGCTTAACCTCAAAATCTTTTTTAACTGCAATTTCTATATAAATGGGACAGT 2240  
QY 2321 ATGCTCTTTAC 2331  
Db 2241 ATGCTCTTTAC 2251

## RESULT 4

US-09-724-676-12752  
; Sequence 12752, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12752  
; LENGTH: 1943  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-12752

Query Match 78.7%; Score 1833.4; DB 5; Length 1943;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1845; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 486 GTGGGATGAAACAAAGTGAAGATCGACCCCGAGTCATATAAATATGATCTTTATATGCC 545  
Db 51 GTGGGATGAAACAAAGTGAAGATCGACCCCGAGTCATATAAATATGATCTTTATATGCC 110  
QY 546 ATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTTGCACGACGAGCTGACACCACTCCC 605  
Db 111 ATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTTGCACGACGAGCTGACACCACTCCC 170  
QY 606 AGCGTGGTGGTCCATGTATGTGAGATGATCAAGAGAGGAGGAAAGAGAGAAATGGAA 665  
Db 171 AGCGTGGTGGTCCATGTATGTGAGATGATCAAGAGAGGAGGAAAGAGAGAAATGGAA 230  
QY 666 AGAATGAGGAGACCTAAAGCCAAATAATTCAGACCCAGGCGCGAGTACACCGCGATC 725  
Db 231 AGAATGAGGAGACCTAAAGCCAAATAATTCAGACCCAGGCGCGAGTACACCGCGATC 290  
QY 726 CACCTCAGCTGAATCTGGCAACGCGAGGAGCGCATTCCTCAATCATCTCACGGGAGGAT 785  
Db 291 CACCTCAGCTGAATCTGGCAACGCGAGGAGCGCATTCCTCAATCATCTCACGGGAGGAT 350  
QY 786 CTTTCTACTGTGGAGTGGCTGCTCAACGCTTTCTTGGAGGTGGCAGCGGAGATCGGGGTG 845  
Db 351 CTTTCTACTGTGGAGTGGCTGCTCAACGCTTTCTTGGAGGTGGCAGCGGAGATCGGGGTG 410  
QY 846 GCAGAAATCCCAAGTTTCATGTTTGTCTCAGAAAGAAATCAAGGCCGTGTCCCTTGTCTTAAT 905  
Db 411 GCAGAAATCCCAAGTTTCATGTTTGTCTCAGAAAGAAATCAAGGCCGTGTCCCTTGTCTTAAT 470  
QY 906 GCTGCACACAGCTTACTGTTTCATGGCACCCGGGAATGACTTGGGCCAATCACTGAGTTTG 965  
Db 471 GCTGCACACAGCTTACTGTTTCATGGCACCCGGGAATGACTTGGGCCAATCACTGAGTTTG 530  
QY 966 TGGTGTATCGCAACAGGACATTTGGGACTGTCTTGAGAAACAGATATGATGATGTTTGG 1025  
Db 531 TGGTGTATCGCAACAGGACATTTGGGACTGTCTTGAGAAACAGATATGATGATGTTTGG 590  
QY 1026 TACTTGTCTTTTCTGTAGGTTCTGTCTGTGCCAAGGGCAGGTTCATCAGTGAAGCTCAG 1085  
Db 591 TACTTGTCTTTTCTGTAGGTTCTGTCTGTGCCAAGGGCAGGTTCATCAGTGAAGCTCAG 650  
QY 1086 GAGAGAGCTTCTGTTTCTAAGTGGCTGACGGGGCACTCTCTACTGTGGAGAGAGGT 1145  
Db 651 GAGAGAGCTTCTGTTTCTAAGTGGCTGACGGGGCACTCTCTACTGTGGAGAGAGGT 710  
QY 1146 ACCACAGGAAGCCGCTTAGTGACAGAGGTTGTGAAAAACAGACGCAATGCAATGTCGAAA 1205  
Db 711 ACCACAGGAAGCCGCTTAGTGACAGAGGTTGTGAAAAACAGACGCAATGCAATGTCGAAA 770  
QY 1206 TTGTAGGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGATGTATATTAATGATTTA 1265

Db 771 TTGTAGCGTTTCCTTTCCCTCATGTTCTCATGTTTGATGATATTA 830  
 Qy 1266 CAAGACTAACCTTTGTTGTTATTAAGTTACACCGTTGTTTTCATCTTTGGAA 1325  
 Db 831 CAAGACTAACCTTTGTTGTTATTAAGTTACACCGTTGTTTTCATCTTTGGAA 890  
 Qy 1326 GCCAGGAAAGCGTTTGGAAAGGATGACCTTTCCGAGATCTCGGATCTCGA 1385  
 Db 891 GCCAGGAAAGCGTTTGGAAAGGATGACCTTTCCGAGATCTCGGATCTCGA 950  
 Qy 1386 GCAGAGACCTGTTGGGAACTCTTCTGGAATGATGATGATGATGATGATGAT 1445  
 Db 951 GCAGAGACCTGTTGGGAACTCTTCTGGAATGATGATGATGATGATGATGAT 1010  
 Qy 1446 GCAGAGACCTGTTGGGAACTCTTCTGGAATGATGATGATGATGATGATGAT 1505  
 Db 1011 GCAGAGACCTGTTGGGAACTCTTCTGGAATGATGATGATGATGATGATGAT 1070  
 Qy 1506 GCTTATACATTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1565  
 Db 1071 GCTTATACATTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1130  
 Qy 1566 GTTTGGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625  
 Db 1131 GTTTGGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1190  
 Qy 1626 ACATGATATAGCCAAAACTTTTGG-AAAAGTTGGTTGATGATGATGATGAT 1684  
 Db 1191 ACATGATATAGCCAAAACTTTTGG-AAAAGTTGGTTGATGATGATGATGAT 1250  
 Qy 1685 ACCCCAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744  
 Db 1251 ACCCCAGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1310  
 Qy 1745 TTGTACACTTGGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1804  
 Db 1311 TTGTACACTTGGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1370  
 Qy 1805 ATCATATGTTGAGGAAAGCTTTCCATGTTCAACAGGATGATGATGATGATGAT 1864  
 Db 1371 ATCATATGTTGAGGAAAGCTTTCCATGTTCAACAGGATGATGATGATGATGAT 1430  
 Qy 1865 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1924  
 Db 1431 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490  
 Qy 1925 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1984  
 Db 1491 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1550  
 Qy 1985 CTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2044  
 Db 1551 CTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1610  
 Qy 2045 ATACTTAATGCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2104  
 Db 1611 ATACTTAATGCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1670  
 Qy 2105 AGAAGCAATATGTTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2164  
 Db 1671 AGAAGCAATATGTTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1730  
 Qy 2165 TTCTCAAACTGTTTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2224  
 Db 1731 TTCTCAAACTGTTTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 1790  
 Qy 2225 TCGCTGCTCAAGTGTCTTGTAAAGTAAATGATGATGATGATGATGATGATGATGAT 2284  
 Db 1791 TCGCTGCTCAAGTGTCTTGTAAAGTAAATGATGATGATGATGATGATGATGAT 1850  
 Qy 2285 TTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2331  
 Db 1851 TTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1897

RESULT 5  
 US-09-724-676A-12752  
 / Sequence 12752, Application US/09724676A  
 / GENERAL INFORMATION:  
 / APPLICANT: Comugen LTD  
 / TITLE OF INVENTION: Variants of alternative splicing  
 / FILE REFERENCE: 129181.4 Comugen  
 / CURRENT APPLICATION NUMBER: US/09/724,676A  
 / NUMBER OF SEQ ID NOS: 97222  
 / SOFTWARE: PatentIn version 3.2  
 / SEQ ID NO 12752  
 / LENGTH: 1943  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / US-09-724-676A-12752  
 Query Match 78.7%; Score 1833.4; DB 5; Length 1943;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1845; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 486 GTGGATGGAACAAGTGAAGATGGAAGCCAGTCAATTAATGATGATGATGATGATGAT 545  
 Db 51 GTGGATGGAACAAGTGAAGATGGAAGCCAGTCAATTAATGATGATGATGATGATGAT 110  
 Qy 546 ATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTTGACGACGACGACGACGACGACGAC 605  
 Db 111 ATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTTGACGACGACGACGACGACGACGAC 170  
 Qy 606 AGCGTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665  
 Db 171 AGCGTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 230  
 Qy 666 AGAATGAGGAGACCTAAGCCAAAAATATCAGACGACGACGACGACGACGACGACGAC 725  
 Db 231 AGAATGAGGAGACCTAAGCCAAAAATATCAGACGACGACGACGACGACGACGACGAC 290  
 Qy 726 CACCTCAGCTGAATGCGACGCGACGACGACGACGACGACGACGACGACGACGACGAC 785  
 Db 291 CACCTCAGCTGAATGCGACGCGACGACGACGACGACGACGACGACGACGACGACGAC 350  
 Qy 786 CTTTACTGTTGAGTGTGCTGACGACGACGACGACGACGACGACGACGACGACGACGAC 845  
 Db 351 CTTTACTGTTGAGTGTGCTGACGACGACGACGACGACGACGACGACGACGACGACGAC 410  
 Qy 846 GCAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGTCCTTGTCTAAT 905  
 Db 411 GCAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGTCCTTGTCTAAT 470  
 Qy 906 GCTGACACACAGTACTGTTTCAATGCGACCCGGGAAATGACTGGGCAATCACTAGTTG 965  
 Db 471 GCTGACACACAGTACTGTTTCAATGCGACCCGGGAAATGACTGGGCAATCACTAGTTG 530  
 Qy 966 TGGTATGCGCAAGAGACATTTGGAGCTGTCTTGAGAAAACAGATATGATGATGATGAT 1025  
 Db 531 TGGTATGCGCAAGAGACATTTGGAGCTGTCTTGAGAAAACAGATATGATGATGATGAT 590  
 Qy 1026 TACTTGTCTTTCGTGAGTTCGTCTGTGCAAGGGCAGGTTGATCAGTCACTCAG 1085  
 Db 591 TACTTGTCTTTCGTGAGTTCGTCTGTGCAAGGGCAGGTTGATCAGTCACTCAG 650  
 Qy 1086 GAGAGAGCTTCTGTTTCAATGAGCTGACAGGCGCACTCTCTAGTGAAGAGGT 1145  
 Db 651 GAGAGAGCTTCTGTTTCAATGAGCTGACAGGCGCACTCTCTAGTGAAGAGGT 710  
 Qy 1146 ACCAAGGAAGCGCTAGTGCAGAGAGGTGTGAAGAACAGCAGCAATGCAATGTGAAA 1205  
 Db 711 ACCAAGGAAGCGCTAGTGCAGAGAGGTGTGAAGAACAGCAGCAATGCAATGTGAAA 770  
 Qy 1206 TTGTAGGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGATGATTA 1265

```
Db 771' TTGTAGCGTTTCCTTTCTTCCCTCATGTTCTCATGTTTGTGCAATGATATTAATTA 830
Qy 1266 CAAGACTAACCTTTGTTTCATATAAAGTTACACCGTTGTTGTTTATACATCTTTGGGAA 1325
Db 831 CAAGACTAACCTTTGTTTCATATAAAGTTACACCGTTGTTGTTTATACATCTTTGGGAA 890
Qy 1326 GCAGGAAGCGTTTGGAAACGATACACCTTTCCAGATTCTCGGATCTCGACTCTTT 1385
Db 891 GCAGGAAGCGTTTGGAAACGATACACCTTTCCAGATTCTCGGATCTCGACTCTTT 950
Qy 1386 GCAACAGCACTTGTCTGGCGAACTCTTCTCGAATGCATTCACCTCAGCATCCCCAAGCGT 1445
Db 951 GCAACAGCACTTGTCTGGCGAACTCTTCTCGAATGCATTCACCTCAGCATCCCCAAGCGT 1010
Qy 1446 GCAACGTTGTAAGTGTCTTTCGAAAGAGTTGATCTGAAATTCCTCTGTGAATTTA 1505
Db 1011 GCAACGTTGTAAGTGTCTTTCGAAAGAGTTGATCTGAAATTCCTCTGTGAATTTA 1070
Qy 1506 GCTTATACAAATTCAGAGTAATAGCACTTCTGCACTGCACTTTTCTGTTGTTTCTAT 1625
Db 1071 GCTTATACAAATTCAGAGTAATAGCACTTCTGCACTGCACTTTTCTGTTGTTTCTAT 1565
Qy 1626 GCTTATAGTTTGGATCGGACCTGCTGTTCTGTTGTTTCTGTTGTTTCTAT 1684
Db 1131 GCTTATAGTTTGGATCGGACCTGCTGTTCTGTTGTTTCTGTTGTTTCTAT 1190
Qy 1685 ACATGAATCATAGCAAAACCTTTTGG- AAACGTTGTTGTTGAGATAGTTGTTCTTT 1684
Db 1191 ACATGAATCATAGCAAAACCTTTTGGAAACCTGTTGTTGAGATAGTTGTTCTTT 1250
Qy 1685 ACCCAGGAGACATCAAGATACACTTGTAAATAAGCTCATAGCATATATTCATACCTG 1744
Db 1251 ACCCAGGAGACATCAAGATACACTTGTAAATAAGCTCATAGCATATATTCATACCTG 1310
Qy 1745 TTGTACACTTGGGTGAAAGTATGCGAGTGGAGACTTAAGATGATTAACCTACCTGTA 1804
Db 1311 TTGTACACTTGGGTGAAAGTATGCGAGTGGAGACTTAAGATGATTAACCTACCTGTA 1370
Qy 1805 ATCATATGTTGTAGGAAAGCTGTTCCATGCTTAACAGGACTTGAATCAAGAGATGTC 1864
Db 1371 ATCATATGTTGTAGGAAAGCTGTTCCATGCTTAACAGGACTTGAATCAAGAGATGTC 1430
Qy 1865 AAGTGGATAGATCTGTGGCGATATGAGAGGATGCACTGCTTCCCACTTCAATCC 1924
Db 1431 AAGTGGATAGATCTGTGGCGATATGAGAGGATGCACTGCTTCCCACTTCAATCC 1490
Qy 1925 TGATGGAATGTTATAGTTAAATTTGATTTTCTGTTGTTTCTGTTGTTTCTGTTGTT 1984
Db 1491 TGATGGAATGTTATAGTTAAATTTGATTTGTAATTTTCTGTTGTTTCTGTTGTT 1550
Qy 1985 CTGGTAAATAGGTTATATATTTTGGCTTACAAATACCGTAACAATGTTGTTCAATTTGAA 2044
Db 1551 CTGGTAAATAGGTTATATATTTTGGCTTACAAATACCGTAACAATGTTGTTCAATTTGAA 1610
Qy 2045 ATACTTAATGCCAAGTAAACATGCTTTGGAATTTGGAAGATGTTTATTTCTTTG 2104
Db 1611 ATACTTAATGCCAAGTAAACATGCTTTGGAATTTGGAAGATGTTTATTTCTTTG 1670
Qy 2105 AGAAGCAAAATATGTTTGCATTAATGCTTTGATTTGTTCAATCAAGAAATGATTTGAACG 2164
Db 1671 AGAAGCAAAATATGTTTGCATTAATGCTTTGATTTGTTCAATCAAGAAATGATTTGAACG 1730
Qy 2165 TTCTCAACCCCTGTTTACCGGATCTTGGTAAGAGGAGCGCGTTTGGGAGAGACCAATTCGA 2224
Db 1731 TTCTCAACCCCTGTTTACCGGATCTTGGTAAGAGGAGCGCGTTTGGGAGAGACCAATTCGA 1790
Qy 2225 TCGCTGTCCAAGTGTCTTGTGTAAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATAC 2284
Db 1791 TCGCTGTCCAAGTGTCTTGTGTAAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATAT 1850
Qy 2285 TTTTAACTGCAATCTATATAAATGGGACAGATAGTCTCTTAC 2331
Db 1851 TTTTAACTGCAATCTATATAAATGGGACAGATAGTCTCTTAC 1897
```

## RESULT 6

```
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)-(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47)-(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756
```

```
Query Match 67.6%; Score 1575.4; DB 5; Length 2050;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 2; Indels 367; Gaps 2;
```

```
Qy 1 TTTTCTTTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAGCACCGCTCAAGA 60
Db 39 TTTTCTTTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAGCACCGCTCAAGA 98
Qy 61 AGGAACCTACAGCCTCTTGGAAAGGAATCTCACTAGGGGCTTGACTGCTGGGCTCTAG 120
Db 99 AGGAACCTACAGCCTCTTGGAAAGGAATCTCACTAGGGGCTTGACTGCTGGGCTCTAG 158
Qy 121 CGCTTTCACCTGTAAGAAAGCAAGATGCAATTTAGAAACTTTTAACCTAGCTCCC 180
Db 159 CGCTTTCACCTGTAAGAAAGCAAGATGCAATTTAGAAACTTTTAACCTAGCTCCC 218
Qy 181 TGATTGCTGTGTGGCAAAACAGTATATCTTTCAGCGAAAGTGAACCCAGGCCAAATTTG 240
Db 219 TGATTGCTGTGTGGCAAAACAGTATATCTTTCAGCGAAAGTGAACCCAGGCCAAATTTG 278
Qy 241 AGTCCCTCTTTAGGACGCTATGACAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAC 300
Db 279 AGTCCCTCTTTAGGACGCTATGACAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAC 338
Qy 301 GAGTCAGAAATAAACTTCAGCAACCCCTTCTCCGAGGAGATGCCAGGCTCCAGCTGCATA 360
Db 339 GAGTCAGAAATAAACTTCAGCAACCCCTTCTCCGAGGAGATGCCAGGCTCCAGCTGCATA 398
Qy 361 AGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 420
Db 399 AGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 458
Qy 421 GCTCAGACCTGGCTCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTC 480
Db 459 GCTCAGACCTGGCTCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTC 518
Qy 481 CGCCAGTGGGATGGAAACAAGTGGAAAGATCCAGACCAAGCAGTTTCTGATCTCCCTCCCGCTCTC 540
Db 519 CGCCAGTGGGATGGAAACAAGTGGAAAGATCCAGACCAAGCAGTTTCTGATCTCCCTCCCGCTCTC 578
Qy 541 ATGCCATCTCCAAGTGGGGCCAGGGGAAAGTATGAATTCACGAGGAGCTGACACCA 600
Db 579 ATGCCATCTCCAAGTGGGGCCAGGGGAAAGTATGAATTCACGAGGAGCTGACACCA 638
Qy 601 CTCCCAGCGTGGTCTCATGTATGTGAGAGTGATCAAGAGAGGAGGAAGAGGAGAA 660
```



Db 39\* TTTTNTTCCAGGAGTGGGGCTGGCCCTTACTGCTTTTATAAGACACAGCTCAAGA 98  
Qy 61 AGGAACCTACAGCCTCTTTGGAAAGAAATCTCACTAGGGGCTTGACTGCTGGGTCTGTAG 120  
Db 99 AGGAACCTACAGCCTCTTTGGAAAGAAATCTCACTAGGGGCTTGACTGCTGGGTCTGTAG 158  
Qy 121 CGCTTTCCTCTAAGAAAGCAGATGCATTTTGAACCTTTAACTAAGTTTGTAGTCCC 180  
Db 159 CGCTTTCCTCTAAGAAAGCAGATGCATTTTGAACCTTTAACTAAGTTTGTAGTCCC 218  
Qy 181 TGAATGCTGTGTGGCAACAGTGTATCTTTCAGCAAGAGTGAACAGGCGCAAAATTTG 240  
Db 219 TGAATGCTGTGTGGCAACAGTGTATCTTTCAGCAAGAGTGAACAGGCGCAAAATTTG 278  
Qy 241 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAC 300  
Db 279 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAC 338  
Qy 301 GAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGTGCATA 360  
Db 339 GAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGTGCATA 398  
Qy 361 AGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 420  
Db 399 AGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 458  
Qy 421 GCTCAGACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCCGCCTCTC 480  
Db 459 GCTCAGACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCCGCCTCTC 518  
Qy 481 CGCCAGTGGGATGAAACAAAGTGGAGATCGGACCCCAAGTATGATCTCCCTCCCGCCTCTC 540  
Db 519 CGCCAGTGGGATGAAACAAAGTGGAGATCGGACCCCAAGTATGATCTCCCTCCCGCCTCTC 578  
Qy 541 ATGCCATCTCCAGCTGGGSCCAGGGGAAAGTATGAATTCAGCAGCAGCTGACACCA 600  
Db 579 ATGCCATCTCCAGCTGGGSCCAGGGGAAAGTATGAATTCAGCAGCAGCTGACACCA 638  
Qy 601 CTCCAGCGTGGTCCATGTATGTAGAGTGATCAAGAGAAAGGAGGAGGAA 660  
Db 639 CTCCAGCGTGGTCCATGTATGTAGAGTGATCAAGAGAAAGGAGGAGGAA 698  
Qy 661 TGGAAAGAAATGAGGAGACCTTAAGCCAAAATATCCAGACCCAGGAGCGGAGTACAGC 720  
Db 699 TGGAAAGAAATGAGGAGACCTTAAGCCAAAATATCCAGACCCAGGAGCGGAGTACAGC 758  
Qy 721 CGATCCACCTCAGCTGAATGCGCACGCGACGAGGACGCAATCCAAATCATACTCAGCGA 780  
Db 759 CGATCCACCTCAGCTGAATGCGCACGCGACGAGGACGCAATCCAAATCATACTCAGCGA 818  
Qy 781 GGAATCTTTTACTGTGGAGTGGCTGGTCACGACTTCTTCGGAGTGGCAGCGAGATCG 840  
Db 819 GGA----- 822  
Qy 841 GGGTGGCAGAAATCCAGTTTCTGCTCAGAGAGAAATCAAGCGGCTGCTCCCTTGT 900  
Db 823 ----- 822  
Qy 901 CTAACTGCTGCACCGATTACTGTTCAATGGCACCCGGGAAATGCTTGGGCCAATCACTGA 960  
Db 823 ----- 822  
Qy 961 GTTTGTGTGATGCGACAGGACATTTGGGACTGTCTTGAGAAACAGATAATGATGTG 1020  
Db 823 ----- 822  
Qy 1021 TTTTGTACTTGTCTTTCTGTAGTGTCTGTCTGCGCAAGGCGAGGTGATCAGTGAG 1080  
Db 823 ----- 822  
Qy 1081 CTCAGGAGAGACTTCCCTGTTTCTAAGTGGCTGAGGGGCACTCTCTACTGTGTAGGAA 1140  
Db 823 ----- 822

Qy 1141 GAGGTACCA CAGGAAGCCGCTAGTGCAGAGAGTTGTGAAACAGCAGCAATGCAATGT 1200  
Db 823 ----- AGGAAGCCGCTAGTGCAGAGAGTTGTGAAACAGCAGCAATGCAATGT 872  
Qy 1201 GGAATTTGTAGCGTTTCCCTTCTCCCTCATCTTCTCATGTTTGTGCGATGATATTAATCTG 1260  
Db 873 GGAATTTGTAGCGTTTCCCTTCTCCCTCATCTTCTCATGTTTGTGCGATGATATTAATCTG 932  
Qy 1261 ATTTCAAGACTAACTTTGTTTGTATATAAAGTTTACACCGTTGTTTGTATACATCTTT 1320  
Db 933 ATTTCAAGACTAACTTTGTTTGTATATAAAGTTTACACCGTTGTTTGTATACATCTTT 992  
Qy 1321 GSGAGCCAGGAAAGCGTTTGGAAAACGATACACCTTTCCAGATCTCGGATCTCGAC 1380  
Db 993 GSGAGCCAGGAAAGCGTTTGGAAAACGATACACCTTTCCAGATCTCGGATCTCGAC 1052  
Qy 1381 TCTTTGCAACAGCACTTGTCTCGGAACTCTTCTCTGGAATGCAATTCATCTCAGCATCCCA 1440  
Db 1053 TCTTTGCAACAGCACTTGTCTCGGAACTCTTCTCTGGAATGCAATTCATCTCAGCATCCCA 1112  
Qy 1441 ACCGTGCAACGTGTAACTTTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGA 1500  
Db 1113 ACCGTGCAACGTGTAACTTTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGA 1172  
Qy 1501 ATTTAGCTTATACATTCAGAGATAGCAGTTTCTCTGCCAACTTTTGTAGTGGGTGAGAAA 1560  
Db 1173 ATTTAGCTTATACATTCAGAGATAGCAGTTTCTCTGCCAACTTTTGTAGTGGGTGAGAAA 1232  
Qy 1561 TTTTAGTTTAGGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTTTATGTGGTGT 1620  
Db 1233 TTTTAGTTTAGGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTTTATGTGGTGT 1292  
Qy 1621 TCTATACATGAATCATGCCAAAACCTTTTGTGG-AAACTGTTGGTTGAGATAGTTGGT 1679  
Db 1293 TCTATACATGAATCATGCCAAAACCTTTTGTGGAAAACCTGTTGGTTGAGATAGTTGGT 1352  
Qy 1680 CTTTACCCCGCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCAT 1739  
Db 1353 CTTTACCCCGCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCAT 1412  
Qy 1740 ACTGTTGTACACTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACC 1799  
Db 1413 ACTGTTGTACACTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACC 1472  
Qy 1800 TGTGAATCATATGTTGTAGGAAAAGCTTGTCCCATGTCTTAACAGGACTTGAATTCAAAGC 1859  
Db 1473 TGTGAATCATATGTTGTAGGAAAAGCTTGTCCCATGTCTTAACAGGACTTGAATTCAAAGC 1532  
Qy 1860 ATGTCAAAGTGGATAGTACTGTTGGCGATATGAGGGGATGAGTGCCTTCCCAATTC 1919  
Db 1533 ATGTCAAAGTGGATAGTACTGTTGGCGATATGAGGGGATGAGTGCCTTCCCAATTC 1592  
Qy 1920 ATTCTGATGGAATGTTTACTAGGTTAACTTTGTAATTTTTTCTAGTGTGAATGTG 1979  
Db 1593 ATTCTGATGGAATGTTTACTAGGTTAACTTTGTAATTTTTTCTAGTGTGAATGTG 1852  
Qy 1980 TATGCTCTGTAATAGTATATATTTTGGCCCTTCAATACCGTTAACTGTTTGTCAAT 2039  
Db 1653 TATGCTCTGTAATAGTATATATTTTGGCCCTTCAATACCGTTAACTGTTTGTCAAT 1712  
Qy 2040 TTGAATTAATTAATGCAAGTAACTATGCTTTGGAAATTTGGAAGATGTTTAT 2099  
Db 1713 TTGAATTAATTAATGCAAGTAACTATGCTTTGGAAATTTGGAAGATGTTTAT 1772  
Qy 2100 CTTTGAAGCAAAATATGTTTGCATTAATATGCTTTGATTTTCAATCAAGAAATGAT 2159  
Db 1773 CTTTGAAGCAAAATATGTTTGCATTAATATGCTTTGATTTTCAATCAAGAAATGAT 1832  
Qy 2160 GAACGTTCTCAAAACCTGTTTACGTTACTTGGTAAAGGGAGCGGTTTGGGAGAGACCA 2219  
Db 1833 GAACGTTCTCAAAACCTGTTTACGTTACTTGGTAAAGGGAGCGGTTTGGGAGAGACCA 1892

QY 2220 TTGCATCGCTGTCACAGTGTCTTGTAGTCTTTAACTGAGAGGCTAACCTCA 2279  
 DB 1893 TTGCATCGCTGTCACAGTGTCTTGTAGTCTTTAACTGAGAGGCTAACCTCA 1952  
 QY 2280 AATACCTTTTAACTGATCTCATATATATATGAGGACAGATGTCTCTTAC 2331  
 DB 1953 AATATTTTAACTGATCTCATATATATATGAGGACAGATGTCTCTTAC 2004

## RESULT 8

US-09-724-676-12748

Sequence 12748, Application US/09724676

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12748

LENGTH: 1931

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (11)..(11)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (18)..(18)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (23)..(23)

OTHER INFORMATION: n is a,c,g, or t

US-09-724-676-12748

Query Match 63.2%; Score 1472.8; DB 5; Length 1931;  
 Best Local Similarity 83.4%; Pred. No. 0;  
 Matches 1861; Conservative 0; Mismatches 3; Indels 367; Gaps 2;

QY 102 TGACTGCTGCTGTGAGGCTTTCACCTGTAAGAAAGCAAGATGATTTAGAACTTT 161  
 DB 21 TACTGCTGCTGTGAGGCTTTCACCTGTAAGAAAGCAAGATGATTTAGAACTTT 80  
 QY 162 AACTACAGTTTACCTCCCTGATTTGCTGTGGCAAAACAGTGTATCTCAGGGAAGT 221  
 DB 81 AACTACAGTTTACCTCCCTGATTTGCTGTGGCAAAACAGTGTATCTCAGGGAAGT 140  
 QY 222 GAAACGAGGGGCAAAATTTAGATCCCTCTTATGAGAGTATGACAGACATCACCTTTCAG 281  
 DB 141 GAAACGAGGGGCAAAATTTAGATCCCTCTTATGAGAGTATGACAGACATCACCTTTCAG 200  
 QY 282 TATTTAAGAGCTTCAACGAGTCAAGATTAACCTTCAAGCAACCTTCTCCGACGAGAT 341  
 DB 201 TATTTAAGAGCTTCAACGAGTCAAGATTAACCTTCAAGCAACCTTCTCCGACGAGAT 260  
 QY 342 GCCAGGCTCCAGCTGATTAAGTGAATTTCTGGGAAAGGAAATGAATATATTTTGGT 401  
 DB 261 GCCAGGCTCCAGCTGATTAAGTGAATTTCTGGGAAAGGAAATGAATATATTTTGGT 320  
 QY 402 CAGACCTTACATAGAGAGTCAACCTGCTCCGCAAAATCCAGACAGAGCTTTCTG 461  
 DB 321 CAGACCTTACATAGAGAGTCAACCTGCTCCGCAAAATCCAGACAGAGCTTTCTG 380  
 QY 462 ATCTCCCTCCGCTCTCCGCAAGTGGATGGAAACAAGTGAAGATGCAACCCGAGTC 521  
 DB 381 ATCTCCCTCCGCTCTCCGCAAGTGGATGGAAACAAGTGAAGATGCAACCCGAGTC 440  
 QY 522 ATAACTATGATCTTTATAGCCATCTCCAAAGTGGGGCCAGGGGAAATGATGAATTTG 581  
 DB 441 ATAACTATGATCTTTATAGCCATCTCCAAAGTGGGGCCAGGGGAAATGATGAATTTG 500

QY 582 CACGACGAGCTGACACCACTCCCAAGCTGTGTGCTCATGTATGTAGAGTATCAAGAG 641  
 DB 501 CACGACGAGCTGACACCACTCCCAAGCTGTGTGCTCATGTATGTAGAGTATCAAGAG 560  
 QY 642 AAGAGAGAAAGAGGAAATGAGAAATGAGAGAGCTTAAGCAAAATTTATCCAGACC 701  
 DB 561 AAGAGAGAAAGAGGAAATGAGAAATGAGAGAGCTTAAGCAAAATTTATCCAGACC 620  
 QY 702 AGAGAGCCGAGTACAGCCGATCCACCTCAGCTGAACTGGCAAGGACGAGAGCATT 761  
 DB 621 AGAGAGCCGAGTACAGCCGATCCACCTCAGCTGAACTGGCAAGGACGAGAGCATT 680  
 QY 762 CCAATATATCTACGCGGAGGATCTTTTACTGTGAGAGGTGGTGTGACGACTTCTTG 821  
 DB 681 CCAATATATCTACGCGGAGGAGAA----- 703  
 QY 822 GAGGTGCAAGCCGAGATCGGGGTGGCAAAATCCAGTTCATGTGTCTCAGAGAGATC 881  
 DB 704 ----- 703  
 QY 882 AAGCCGATCCCTGTGTCTATGTCTGCAACACAGTTACTGTTCATGGAACCGGGAAT 941  
 DB 704 ----- 703  
 QY 942 GACTTGGCCATCACTGAGTTGTGTGATCGCAAGAGACATTTGGAGCTGTCTGAG 1001  
 DB 704 ----- 703  
 QY 1002 AAAACAGATATATATAGTGTGTGTGATCTTTCTGTGTAGTGTCTGTGTGCCAA 1061  
 DB 704 ----- 703  
 QY 1062 GGGCAGTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTGTCTAAGTGGCTGAGGGCC 1121  
 DB 704 ----- 703  
 QY 1122 CACTCTTACTGTAGAGAGAGTACACAGAGAGCCGCTAGTGGAGAGGTTGGAA 1181  
 DB 704 -----AGAGAGCCGCTAGTGGAGAGGTTGGAA 734  
 QY 1182 AACAGAGAGATCAATGTGAAATTTGAGCGTTTCTTCTTCTCCTCATGTCTCATGT 1241  
 DB 735 AACAGAGAGATCAATGTGAAATTTGAGCGTTTCTTCTTCTCCTCATGTCTCATGT 794  
 QY 1242 TTGTGATGTATATTTACTGATTTTACAGACTTACCTTGTGTGTATTAAGTTACACG 1301  
 DB 795 TTGTGATGTATATTTACTGATTTTACAGACTTACCTTGTGTGTATTAAGTTACACG 854  
 QY 1302 TTGTGTTTTACATCTTTTGGGAAGCCAGGAAAGGTTTGGAAACGTATCACCTTCC 1361  
 DB 855 TTGTGTGTTTTACATCTTTTGGGAAGCCAGGAAAGGTTTGGAAACGTATCACCTTCC 914  
 QY 1362 AGATTTCTCGATTTCTGACTCTTTCGCAACAGCATCTTTCGGAAGTCTTCTGGAATG 1421  
 DB 915 AGATTTCTCGATTTCTGACTCTTTCGCAACAGCATCTTTCGGAAGTCTTCTGGAATG 974  
 QY 1422 CATTCATCAGATCCCAACCGTGCAGCTTAACCTTGTGTGTGTGCAAAAGAGTGA 1481  
 DB 975 CATTCATCAGATCCCAACCGTGCAGCTTAACCTTGTGTGTGTGCAAAAGAGTGA 1034  
 QY 1482 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGATAGCACTTCACTGCA 1541  
 DB 1035 TCTGAAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGATAGCACTTCACTGCA 1094  
 QY 1542 ACTTTAGTGGTGAAGAAATTTAGTTAGTGTGTGTGAGTGGACCTCAGTTCTGTGG 1601  
 DB 1095 ACTTTAGTGGTGAAGAAATTTAGTTAGTGTGTGTGAGTGGACCTCAGTTCTGTGG 1154  
 QY 1602 TTTCTTTATGTGTGTGTCTTATATGAAATCATAGCCAAATCTTTTGG-AAAAGT 1660  
 DB 1155 TTTCTTTATGTGTGTGTCTTATATGAAATCATAGCCAAATCTTTTGGAAAACTG 1214



```

QY 1661 TTGGTTGAGATAGTTGGTCTTTTACCCCAAGACATCAAGATACACTTGTAAATAAA 1720
Db 1215 TTGGTTGAGATAGTTGGTCTTTTACCCCAAGACATCAAGATACACTTGTAAATAAA 1274
QY 1721 GCTGATAGCATATATTCATACCTGTTGTACACTTGGGTGAAAGATGCGAGTGGGAGAC 1780
Db 1275 GCTGATAGCATATATTCATACCTGTTGTACACTTGGGTGAAAGATGCGAGTGGGAGAC 1334
QY 1781 TAAGATGTATTAACCTACCTGTGAAATCATATGTTGTAGGAAAGCTGTCCCATGTCTAA 1840
Db 1335 TAAGATGTATTAACCTACCTGTGAAATCATATGTTGTAGGAAAGCTGTCCCATGTCTAA 1394
QY 1841 CAGGACTTGAATTCAAAGCATGCAAGTGGATAGTACATCTGTGCGATATGAGAGGAT 1900
Db 1395 CAGGACTTGAATTCAAAGCATGCAAGTGGATAGTACATCTGTGCGATATGAGAGGAT 1454
QY 1901 CGAGTGCCTTTCCCATTCATTCCTGTGAGAAATGTTTATACCTAGGTTAAACATTTGTAAT 1960
Db 1455 CGAGTGCCTTTCCCATTCATTCCTGTGAGAAATGTTTATACCTAGGTTAAACATTTGTAAT 1514
QY 1961 TTTTCTAGTTGTAATGTATGTCTGGTAAATAGGTATTATATTTTGGCCTTACAATAC 2020
Db 1515 TTTTCTAGTTGTAATGTATGTCTGGTAAATAGGTATTATATTTTGGCCTTACAATAC 1574
QY 2021 CGTAACATGTTTGTCAATTTTGAATACTTAATGCCAAGTAACAAATCGATGCTTTGAAA 2080
Db 1575 CGTAACATGTTTGTCAATTTTGAATACTTAATGCCAAGTAACAAATCGATGCTTTGAAA 1634
QY 2081 TTTGAGATAGGTTTATCTTTGAGAAGCAAAATATGTTTGCATTAATGCTTTGATTGT 2140
Db 1635 TTTGAGATAGGTTTATCTTTGAGAAGCAAAATATGTTTGCATTAATGCTTTGATTGT 1594
QY 2141 TCATATCAAGAAATGATGAAAGCTTCTCAACCTGTTTACGGTACTTGGTAAGAGGGA 2200
Db 1695 TCATATCAAGAAATGATGAAAGCTTCTCAACCTGTTTACGGTACTTGGTAAGAGGGA 1754
QY 2201 GCCGGTTGGGAGAGACCATTCATCTGCTCAAGGTTTCTGTTTAAAGTGGCTTTTAA 2260
Db 1755 GCCGGTTGGGAGAGACCATTCATCTGCTCAAGGTTTCTGTTTAAAGTGGCTTTTAA 1814
QY 2261 CTGGAGAGGCTAACCTCAAAATACTTTTTTAACTGATCTATAATAATGGGCACAGT 2320
Db 1815 CTGGAGAGGCTAACCTCAAAATACTTTTTTAACTGATCTATAATAATGGGCACAGT 1874
QY 2321 ATGCTCCTTAC 2331
Db 1875 ATGCTCCTTAC 1885

```

## RESULT 9

```

US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a.c.g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a.c.g, or t
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a.c.g, or t
; US-09-724-676A-12748

```

```

Query Match 63.2%; Score 1472.8; DB 5; Length 1931;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 1861; Conservative 0; Mismatches 3; Indels 367; Gaps 2;

```

```

QY 102 TGACTCGTGGGCTGTAGCGCTTTTCACTGTAAAGAAAGCAAGATGCAATTTTAAAGAACTTT 161
Db 21 TANTCTCGTGGGCTGTAGCGCTTTTCACTGTAAAGAAAGCAAGATGCAATTTTAAAGAACTTT 80
QY 162 AACTACAGTTTTAGCTCCCTGATGCTGTGTGGCAAAACAGTGATATCTTCAGCGAAAGT 221
Db 81 AACTACAGTTTTAGCTCCCTGATGCTGTGTGGCAAAACAGTGATATCTTCAGCGAAAGT 140
QY 222 GAAACAGGCGCCAAATTTTGAAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCTAG 281
Db 141 GAAACAGGCGCCAAATTTTGAAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCTAG 200
QY 282 TATTTTAAGAGCTTCAACAGAGTCAGAAATAAATTTTCAAGAACCCCTTCTCCGACAGCAT 341
Db 201 TATTTTAAGAGCTTCAACAGAGTCAGAAATAAATTTTCAAGAACCCCTTCTCCGACAGCAT 260
QY 342 GCAGGCTCCAGCTGCATAGAAGCTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCT 401
Db 261 GCAGGCTCCAGCTGCATAGAAGCTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCT 320
QY 402 CAGACCTTACATAGGAAGCTCACACCTGCTCCGCAAAATCCAGACAAAGCAAGCTTTCTG 461
Db 321 CAGACCTTACATAGGAAGCTCACACCTGCTCCGCAAAATCCAGACAAAGCAAGCTTTCTG 380
QY 462 ATCTCCCTCCGCTCTCCGCAAGTGGATGGAACAAAGTGGAGATGCGACCCAGTC 521
Db 381 ATCTCCCTCCGCTCTCCGCAAGTGGATGGAACAAAGTGGAGATGCGACCCAGTC 440
QY 522 ATAACTATGATCTCTTATATGCTCATCTCCAGCTGGGGCCAGGGGAAAGTGAATTTG 581
Db 441 ATAACTATGATCTCTTATATGCTCATCTCCAGCTGGGGCCAGGGGAAAGTGAATTTG 500
QY 582 CACGAGGCTGACACACTCCAGCTGGTGTGCTCATGTATGTAGAGATGATCAAGAG 641
Db 501 CACGAGGCTGACACACTCCAGCTGGTGTGCTCATGTATGTAGAGATGATCAAGAG 560
QY 642 AAGGAGGAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAATTTATCCAGACC 701
Db 561 AAGGAGGAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAATTTATCCAGACC 620
QY 702 AGGAGGCGGAGTACACCGGATCCACCTCAGCTGAACTGCGACGCGACGAGGACGCAAT 761
Db 621 AGGAGGCGGAGTACACCGGATCCACCTCAGCTGAACTGCGACGCGACGAGGACGCAAT 680
QY 762 CCAATCATACTCACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTGCTACGACTTCTTCG 821
Db 681 CCAATCATACTCACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTGCTACGACTTCTTCG 703
QY 822 GAGTGGCAGCCGAGATCGGGGTGGCAGAAATCCAGTTTCTGCTCAGAGAGAAATC 881
Db 704 ----- 703
QY 882 AAGGCCGTGTCCTTGTCTTAATGTGTCACACCACTTACTGTTTCATGTCACCCGGGAAT 941
Db 704 ----- 703
QY 942 GACTTGGGCAATCACTAGTGTGTGCTGATCGCAAGGACATTTGGGACTGTCTTGAG 1001
Db 704 ----- 703
QY 1002 AAAACAGATAATGATAGTGTGTTGTACTTGTCTTTCTGTTAGGTCTGCTGTGTCGCAA 1061
Db 704 ----- 703

```

QY 1062 GGGCAGTTGATCACTGAGCTCAGAGAGAGCTTCTGTCTTAAGTGGCTCGAGGGGC 1121  
 Db 704 ----- 703  
 QY 1122 CACTCTACTGTAGAGAGAGTACCAAGAGAGCCGCTTAAGTGCAGAGAGTTGTGA 1181  
 Db 704 ----- AGGAGCCGCGTGTGTGCAAGAGAGTTGTGA 734  
 QY 1183 AACACAGCAATGCAATGTGGAATTTGAGCTTCTCTTCTCTCTCATGTCTCATGT 1241  
 Db 735 AACAGAGCAATGCAATGTGGAATTTGAGCTTCTCTTCTCTCTCATGTCTCATGT 794  
 QY 1242 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1301  
 Db 795 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 854  
 QY 1302 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1361  
 Db 855 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 914  
 QY 1362 AGATTTCTGAGATTTCTGAGCTTCTTGCAGAGCACTTGTGGGAACTCTCTCGAATG 1421  
 Db 915 AGATTTCTGAGATTTCTGAGCTTCTTGCAGAGCACTTGTGGGAACTCTCTCGAATG 974  
 QY 1422 CATTCAGTCAAGTCCCAACCGTGAACGCTGAACCTGTGCTTTTGCAGAGAGTTGA 1481  
 Db 975 CATTCAGTCAAGTCCCAACCGTGAACGCTGAACCTGTGCTTTTGCAGAGAGTTGA 1034  
 QY 1482 TCTGAATTCCTCTGTGAATTTAGCTTAATTAATTAATTAATTAATTAATTAATTA 1541  
 Db 1035 TCTGAATTCCTCTGTGAATTTAGCTTAATTAATTAATTAATTAATTAATTAATTA 1094  
 QY 1542 ACTTTTAACTGGTGAAGAAATTTAGTTAGTGGTGGGATGGGAGCTGAGTTCTGTG 1601  
 Db 1095 ACTTTTAACTGGTGAAGAAATTTAGTTAGTGGTGGGATGGGAGCTGAGTTCTGTG 1154  
 QY 1602 TTCTTTTAACTGGTGGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1660  
 Db 1155 TTCTTTTAACTGGTGGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1214  
 QY 1661 TTGTTGATGATGTTGTTCTTTTACCCCAAGAGAGATTAATTAATTAATTAATTAATTA 1720  
 Db 1215 TTGTTGATGATGTTGTTCTTTTACCCCAAGAGAGATTAATTAATTAATTAATTAATTA 1274  
 QY 1721 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1780  
 Db 1275 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334  
 QY 1781 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1840  
 Db 1335 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1394  
 QY 1841 CAGAGCTTGAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1900  
 Db 1395 CAGAGCTTGAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1454  
 QY 1901 GCAGAGCTTCCCATTTCAATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1960  
 Db 1455 GCAGAGCTTCCCATTTCAATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1514  
 QY 1961 TTTTCTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020  
 Db 1515 TTTTCTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574  
 QY 2021 CGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2080  
 Db 1575 CGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1634  
 QY 2081 TTTGGAAGTGTGTTTATTTCTTGAAGCAAAATGTTGATTAATTAATTAATTAATTAATTA 2140  
 Db 1635 TTTGGAAGTGTGTTTATTTCTTGAAGCAAAATGTTGATTAATTAATTAATTAATTAATTA 1694  
 QY 2141 TCATATCAAGAAATGATGAAAGTTCTCAAAAGCTGTTTACGGTACTTGGTGAAGGGA 2200

Db 1695 TCATATCAAGAAATGATGAAAGTTCTCAAAAGCTGTTTACGGTACTTGGTGAAGGGA 1754  
 QY 2201 GCGGTTTGGGAGAGACCAATTCATGCTGTGCCAGTGTCTTGTGAAGTCTTTTAA 2260  
 Db 1755 GCGGTTTGGGAGAGACCAATTCATGCTGTGCCAGTGTCTTGTGAAGTCTTTTAA 1814  
 QY 2261 CTGAGAGGCTAACTTCAAAATCTTTTAACTGATCTTAATTAATTAATTAATTAATTA 2320  
 Db 1815 CTGAGAGGCTAACTTCAAAATCTTTTAACTGATCTTAATTAATTAATTAATTAATTA 1874  
 QY 2321 ATGCTCTTAC 2331  
 Db 1875 ATGCTCTTAC 1885

RESULT 10  
 US-09-724-676-12741  
 : Sequence 12741, Application US/09724676  
 : GENERAL INFORMATION:  
 : APPLICANT: Compugen LTD  
 : TITLE OF INVENTION: Variants of alternative splicing  
 : FILE REFERENCE: 129181.4 Compugen  
 : CURRENT APPLICATION NUMBER: US/09/724,676  
 : NUMBER OF SEQ ID NOS: 97222  
 : SOFTWARE: PatentIn version 3.2  
 : SEQ ID NO 12741  
 : LENGTH: 1893  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 US-09-724-676-12741

Query Match 58.1%; Score 1354.4; DB 5; Length 1893;  
 Best Local Similarity 80.8%; Pred. No. 0;  
 Matches 1775; Conservative 0; Mismatches 56; Indels 367; Gaps 2;

QY 135 GAAAGCAAGATGATTTTGAAGAACTTTAACTAACAGTTTAACTGCTCCGATGCTGTG 194  
 Db 16 GACTCGAGATGAGAGAGAGTGGACCTGAGAGACCTGCCAGCCGACATGCTGTGAC 75  
 QY 195 GCAACAGTGAATCTTACGCGAAAGTGAACAGAGGCAATTTGATGCTCTTTAGG 254  
 Db 76 CTGAGACCCGCGGTGTGTGAGAGGCTGTGCGGCGCAATTTGAGTCCCTTTAGG 135  
 QY 255 ACGTATGACAGAGATCACTTTTCAATTTTAAAGCTTCAACAGAGTGAATTAAC 314  
 Db 136 ACGTATGACAGAGATCACTTTTCAATTTTAAAGCTTCAACAGAGTGAATTAAC 195  
 QY 315 TTACAGCAACCCCTTCCGAGCAGATGCGAGGCTCCAGCTGCAATTAAGTGAATTTCTG 374  
 Db 196 TTACAGCAACCCCTTCCGAGCAGATGCGAGGCTCCAGCTGCAATTAAGTGAATTTCTG 255  
 QY 375 GGAAGGAATGAATTAATTTTCTGACAGCTTACATGATGAGAGTCAACACTGCTGCT 434  
 Db 196 TTACAGCAACCCCTTCCGAGCAGATGCGAGGCTCCAGCTGCAATTAAGTGAATTTCTG 255  
 QY 435 CCGCAATTCAGACAGAGATTTCTGATCTCCCTCCGCTCCGCGAGTGGAGTGG 494  
 Db 316 CCGCAATTCAGACAGAGATTTCTGATCTCCCTCCGCTCCGCGAGTGGAGTGG 375  
 QY 495 AAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 554  
 Db 376 AAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 435  
 QY 555 CTGGGCTCAGAGGAAAGTGAATTTGACGAGAGTGAACCACTCCAGCGTGTG 614  
 Db 436 CTGGGCTCAGAGGAAAGTGAATTTGACGAGAGTGAACCACTCCAGCGTGTG 495  
 QY 615 GTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674  
 Db 496 GTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555



QY	675	AGACCTAAGCCAAAAATTATCCAGACCAGGAGCGGAGTACACGCCGATCCACCTCAGC	733
Db	556	AGACCTAAGCCAAAAATTATCCAGACCAGGAGCGGAGTACACGCCGATCCACCTCAGC	615
QY	735	TGAACCTGGCAGCGAGGACGCATTTCCAATCATACTACCGGAGGAAATCTTTTACTG	794
Db	616	TGAACCTGGCAGCGAGGACGCATTTCCAATCATACTCACGGAGGAA-----	665
QY	795	TGGAGGTGGCTGGTCCACGACTCTTTCCGAGGTGGCAGCCGAGATCGGGGTGGCAGAAATC	854
Db	666	-----	665
QY	855	CCAGTTTCATGTTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTTTCTTAATGCTGCACAC	914
Db	666	-----	665
QY	915	CAGTTACTGTTTCATGGCACC CGGGAATGACTTTGGGCCAATCACTGAGTTTGTGGTGATCG	974
Db	666	-----	665
QY	975	CACAAGGACATTTGGGACTGCTTTGAGAAAAACAGATAATGATAGTGTGTTGTGTTGTTTC	1034
Db	666	-----	665
QY	1035	TTTTCTGTGTAGTCTTGTCTGTGCGCAAGGGGACGGTTGATCAGTGAGCTCAGGAGAGACT	1094
Db	666	-----	665
QY	1095	TCTGTGTTCTAAGTGGCCCTGCAGGGGCCACTCTCTACTGTTAGGAAGAGTTACCAACAGA	1154
Db	666	-----	AGGA 669
QY	1155	AGCGGCTAGTGCAAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAATTTGTAGCGT	1214
Db	670	AGCGGCTAGTGCAAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAATTTGTAGCGT	729
QY	1215	TTCTCTTCTTCCTCCTCATGTTCTCATGTTTGTGCACTATATTAATTAAGAGACTAA	1274
Db	730	TTCTCTTCTTCCTCCTCATGTTCTCATGTTTGTGCACTATATTAATTAAGAGACTAA	789
QY	1275	CCTTGTTTCGTATATAAAGTTTACACGTTGTGTTTTCATCTTTTGGGAAGCCAGGAAA	1334
Db	790	CCTTGTTTCGTATATAAAGTTTACACGTTGTGTTTTCATCTTTTGGGAAGCCAGGAAA	849
QY	1335	GGTTTGGAAAAACGTATCACTTTCCAGATTCTCGGATTTCTCGAATCTTTTGGCAACAGCA	1394
Db	850	GGTTTGGAAAAACGTATCACTTTCCAGATTCTCGGATTTCTCGAATCTTTTGGCAACAGCA	909
QY	1395	CTTGCTTTCGGGAACTCTTCCTGGAATGCAATCACTCAGCATCCCAACCGTGCACGCTGT	1454
Db	910	CTTGCTTTCGGGAACTCTTCCTGGAATGCAATCACTCAGCATCCCAACCGTGCACGCTGT	969
QY	1455	AACTTGTGCTTTTGGAAAAAGTTCATCTGAAATTCCTCTGTAGAAATTTAGCTTTATACA	1514
Db	970	AACTTGTGCTTTTGGAAAAAGTTCATCTGAAATTCCTCTGTAGAAATTTAGCTTTATACA	1029
QY	1515	ATTCAGAGATAGCAGTTTCACTGCGCACTTTTATGTTGGTGAGAAATTTTATGTTAGTG	1574
Db	1030	ATTCAGAGATAGCAGTTTCACTGCGCACTTTTATGTTGGTGAGAAATTTTATGTTAGTG	1089
QY	1575	TTTGGGATCGGAACTCAGTTTCTGTGTGTTCTTTTATGTGGTGAGAAATTTTATGTTAGTG	1634
Db	1090	TTTGGGATCGGAACTCAGTTTCTGTGTGTTCTTTTATGTGGTGAGAAATTTTATGTTAGTG	1149
QY	1635	ATAGCCAAAAAATTTTTTGG-AAACTGTTGGTTGAGATAGTTGTTCTTATACATGAATC	1693
Db	1150	ATAGCCAAAAAATTTTTTGGAAAACTGTTTGGTTGAGATAGTTGTTCTTATACATGAATC	1209
QY	1694	AGACATCAAGATACACTTGTGAAATAAGCTGATAGCATATTTTCATACCTGTTGTACACT	1753
Db	1210	AGACATCAAGATACACTTGTGAAATAAGCTGATAGCATATTTTCATACCTGTTGTACACT	1269
QY	1754	TGGGTGAAAAAGTATGCGACTGGGAGACTTAAGATGTTTAACTCACTCTGTGGAATCATATGT	1813

Db	1270		TGGGTGAAAGATGTCGACGTGGGAGACTAAGATGTATTAACTACCCTGTGAATCATATGT	1329
Qy	1814		TGTAGGAAAAGCTGTTCCCATGCTCTAACAGCACTTCAAATTCAAAAGCATGTCAAGTGGAATA	1873
Db	1330		TGTAGGAAAAGCTGTTCCCATGCTCTAACAGCACTTCAAATTCAAAAGCATGTCAAGTGGAATA	1389
Qy	1874		GTAGATCTGTGGCGATATGAGAGGGATGCAAGTGCCTTCCCATTCAATCCTGTATGGAAT	1933
Db	1390		GTAGATCTGTGGCGATATGAGAGGGATGCAAGTGCCTTCCCATTCAATCCTGTATGGAAT	1449
Qy	1934		TGTTATCTAGGTTAAACATTGTAAATTTTTTCTAGTGTCTAATGTATGTCTCGTAAAT	1993
Db	1450		TGTTATCTAGGTTAAACATTGTAAATTTTTTCTAGTGTCTAATGTATGTCTCGTAAAT	1509
Qy	1994		AGGTATTATATTTTGGCCTTTACAATACCGTAACAATGTTGTTCATTTTGAATACTTAAAT	2053
Db	1510		AGGTATTATATTTTGGCCTTTACAATACCGTAACAATGTTGTTCATTTTGAATACTTAAAT	1569
Qy	2054		GCCAAGTAACAATGCATGCTTTGGAAATTTGGAAGATGGTTTTATCTTTGAGAAGCAAA	2113
Db	1570		GCCAAGTAACAATGCATGCTTTGGAAATTTGGAAGATGGTTTTATCTTTGAGAAGCAAA	1629
Qy	2114		TATGTTTGCAFTAAATGCTTTGATGTTTCATATCAAGAAATTTGATTGAACGTTCTCAAC	2173
Db	1630		TATGTTTGCAFTAAATGCTTTGATGTTTCATATCAAGAAATTTGATTGAACGTTCTCAAC	1689
Qy	2174		CCTGTTTACGGTACTTGGTAAGGGAGGCCGTTTGGGAGAGACCATTTGCATCGCTGTCC	2233
Db	1690		CCTGTTTACGGTACTTGGTAAGGGAGGCCGTTTGGGAGAGACCATTTGCATCGCTGTCC	1749
Qy	2234		AAGTGTTTCTTGTTAAAGTGCTTTTAACTGGAGAGCGCTAACCTCABAATACTTTTTTAA	2293
Db	1750		AAGTGTTTCTTGTTAAAGTGCTTTTAACTGGAGAGCGCTAACCTCABAATACTTTTTTAA	1809
Qy	2294		CTGCATTCTATAATAATGGGCACAGTATGTCCTTAC	2331
Db	1810		CTGCATTCTATAATAATGGGCACAGTATGTCCTTAC	1847

```

RESULT 11
US-09-724-676A-12741
; Sequence 12741, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741

```

	Query Match	58.1%	Score 1354.4	DB 5	Length 1893
	Best Local Similarity	80.8%	Pred. No. 0		
	Matches 1775	Conservative	0	Mismatches	56
				Indels	367
				Gaps	
Qy	135	GAAGACAGATGCATTTTAGAAGACTTTAACTACAGTTTTAGCTCCCTGATTTGCCGTGTG	194		
Db	16	GACTGCAGATGGAGGAGTGGACCTCGAGACTCTGCCAGCGCCACCATCGCTGTCTCAC	75		
Qy	195	GCAACACAGTGATCTTTGACGGAAAGTGAAACACAGGGCCAAATTTGAGTCCCTCTTTAGG	254		
Db	76	CTGGACCCGGCGGTTCGTGGACGGCGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGG	135		
Qy	255	ACGTATCACAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGAAATAAC	314		
Db	136	ACGTATGCAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACAGGAGTCAGAAATAAC	195		

315 TTCAGCAACCCCTTCTCCGACAGATCCAGGCTCAAGTCATTAAGTGTCTG 374  
 Db 196 TTCAGCAACCCCTTCTCCGACAGATCCAGGCTCAAGTCATTAAGTGTCTG 255  
 Qy 375 GGAAGAAATAGATTAATTTTGTCTCAGACCTTAACATAGAACTCACCTGGCT 434  
 Db 256 GGAAGAAATAGATTAATTTTGTCTCAGACCTTAACATAGAACTCACCTGGCT 315  
 Qy 435 CCGCCAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCCGCGAGTGGATGG 494  
 Db 316 CCGCCAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCCGCGAGTGGATGG 375  
 Qy 495 AAACAGATGAAGATGCGACCCCACTAATATGATCTCTTAATGCCATCTCCAG 554  
 Db 376 AAACAGATGAAGATGCGACCCCACTAATATGATCTCTTAATGCCATCTCCAG 435  
 Qy 555 CTGGGGCCAGAGGAAAGTATGATTCAGACGAGCTGACACCACTCCAGCGTGGTG 614  
 Db 436 CTGGGGCCAGAGGAAAGTATGATTCAGACGAGCTGACACCACTCCAGCGTGGTG 495  
 Qy 615 GTCCATGATGTAGAGTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674  
 Db 496 GTCCATGATGTAGAGTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555  
 Qy 675 AGACCTAAGCCAAATATTCAGACGAGAGGCGGAGTACACGCGCATCCACTCAGC 734  
 Db 556 AGACCTAAGCCAAATATTCAGACGAGAGGCGGAGTACACGCGCATCCACTCAGC 615  
 Qy 735 TGAACTGGACGCGACGAGAGCGCATTCGAAATCATCTCAGCGAGAGAA----- 794  
 Db 616 TGAACTGGACGCGACGAGAGCGCATTCGAAATCATCTCAGCGAGAGAA----- 665  
 Qy 795 TGGAGTGGCTGGTCAAGACTTCTTGAGAGTGGCAGCCGAGATCGGGTGGCAGAAATC 854  
 Db 666 ----- 665  
 Qy 855 CCAATCATGTGCTCAGAGAGATCAAGCCGTGCTCCCTGTCTAATGCTGACAC 914  
 Db 666 ----- 665  
 Qy 915 CAGTACTGTTCAATGCAACCGGAAATGACTTGGCCAACTCACTGATTTGTGTATCG 974  
 Db 666 ----- 665  
 Qy 975 CACAAGCATTGGAGCTGTCTTGAGAAACAGATATGATGTTTGTACTTGTTC 1034  
 Db 666 ----- 665  
 Qy 1035 TTTTCTGTAGGTTCTGTCTGTGCCAAGGCGAGGTGATCAGTGAAGCTCAGAGAGAGCT 1094  
 Db 666 ----- 665  
 Qy 1095 TCTGTCTTAAGTGGCTGCGAGGGGCACTCTTAAGTGAAGAGTACCACAGGA 1154  
 Db 666 ----- 665  
 Qy 1155 AGCCGCTAAGTACAGAGAGTTGTAAGAAACAGACCAATGCAATGTGAATGTAGCGT 1214  
 Db 670 AGCCGCTAAGTACAGAGAGTTGTAAGAAACAGACCAATGCAATGTGAATGTAGCGT 729  
 Qy 1215 TTTCTTTCTCCCTCATGTTCTCATGTTTGTGATGATTAATTTAGTAAAGACTTA 1274  
 Db 730 TTTCTTTCTCCCTCATGTTCTCATGTTTGTGATGATTAATTTAGTAAAGACTTA 789  
 Qy 1275 CTTTGTGTATTAATTAAGTTACACGTTGTTTCAATCTTTTGGAAAGCCAGAAA 1334  
 Db 790 CTTTGTGTATTAATTAAGTTACACGTTGTTTCAATCTTTTGGAAAGCCAGAAA 849  
 Qy 1335 GCGTTTGAAGAAAGTATCACTTCCAGATTTCCGATTTCTGACTTTTGGACAGCA 1394  
 Db 850 GCGTTTGAAGAAAGTATCACTTCCAGATTTCTCGATTTCTGACTTTTGGACAGCA 909  
 Qy 1395 CTTCGTTGGGAACTCTCTCGAATGATTCATCAGATCCCAACCGTGCAGAGTGT 1454

Db 910 CTTGCTTGGGAACTCTCTCTGGAATGATTAATCACTGATCCCAACCGTGCAGAGTGT 969  
 Qy 1455 AACTGTGCTTTTGAAGAAAGATGATCTGAATTTCTGTAGAAATTTAGCTTATACA 1514  
 Db 970 AACTGTGCTTTTGAAGAAAGATGATCTGAATTTCTGTAGAAATTTAGCTTATACA 1029  
 Qy 1515 ATTGAGAAATAGAGTTTCACTGCCAATTTTATGAGGAGAAATTTTATGTTAGGAG 1574  
 Db 1030 ATTGAGAAATAGAGTTTCACTGCCAATTTTATGAGGAGAAATTTTATGTTAGGAG 1089  
 Qy 1575 TTTGGGATGGACCTGAGTTCTGTGTTTCTTTATGAGTGTCTTATAGATGAATC 1634  
 Db 1090 TTTGGGATGGACCTGAGTTCTGTGTTTCTTTATGAGTGTCTTATAGATGAATC 1149  
 Qy 1635 ATTACCAAAATCTTTTGGG-AAAAGTGTGAGATGAGTGTCTTTTACCCACGA 1693  
 Db 1150 ATTACCAAAATCTTTTGGGAAATCTGTGTGAGATGAGTGTCTTTTACCCACGA 1209  
 Qy 1694 AGACATCAAGTATACATTTGTAATTAAGCTGATAGCATATATTCATACCTGTGTACAT 1753  
 Db 1210 AGACATCAAGTATACATTTGTAATTAAGCTGATAGCATATATTCATACCTGTGTACAT 1269  
 Qy 1754 TGGGTGAAGAAAGTATGGCAGTGGAGACTTAAGATGATTAATCTTACCTGTGATATGT 1813  
 Db 1270 TGGGTGAAGAAAGTATGGCAGTGGAGACTTAAGATGATTAATCTTACCTGTGATATGT 1329  
 Qy 1814 TGTAGAAAGCTGTCTCCATGTCTTAACAGACTTGAATTCAGATGTCAGATGATA 1873  
 Db 1330 TGTAGAAAGCTGTCTCCATGTCTTAACAGACTTGAATTCAGATGTCAGATGATA 1389  
 Qy 1874 GTAGATCTGTGGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1933  
 Db 1390 GTAGATCTGTGGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449  
 Qy 1934 TGTATCTAGGTTAACTTTGTAATTTTCTAGTTGTAATGTAATGTCTGTAAAT 1993  
 Db 1450 TGTATCTAGGTTAACTTTGTAATTTTCTAGTTGTAATGTAATGTCTGTAAAT 1509  
 Qy 1994 AGGATTAATTTTGGCTTCAATACCGTAACATGTTTGTCAATTTTGAATTTTGAAT 2053  
 Db 1510 AGGATTAATTTTGGCTTCAATACCGTAACATGTTTGTCAATTTTGAATTTTGAAT 1569  
 Qy 2054 GCCAAGTAACTAGATGCTTGGAAATTTGGAAGAGTGTATTTTCTTGAAGAGCAA 2113  
 Db 1570 GCCAAGTAACTAGATGCTTGGAAATTTGGAAGAGTGTATTTTCTTGAAGAGCAA 1629  
 Qy 2114 TATGTTGATTAATTAATGCTTTGATTTGTAATTCAGAAATTTGTAACGTTTCAAC 2173  
 Db 1630 TATGTTGATTAATTAATGCTTTGATTTGTAATTCAGAAATTTGTAACGTTTCAAC 1689  
 Qy 2174 CCGTTTACGATTAATGCTTTGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233  
 Db 1690 CCGTTTACGATTAATGCTTTGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1749  
 Qy 2234 AAGTGTCTGTGTAAGTCTTTTAACTGAGAGGCTTAACCTCAAAATTTTATTTAA 2293  
 Db 1750 AAGTGTCTGTGTAAGTCTTTTAACTGAGAGGCTTAACCTCAAAATTTTATTTAA 1809  
 Qy 2294 CTGCAATCTAATTAATTAATGGGCACTATGCTCTTAC 2331  
 Db 1810 CTGCAATCTAATTAATTAATGGGCACTATGCTCTTAC 1847

RESULT 12  
 US-09-724-676-12745  
 ; Sequence 12745, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12745  
; LENGTH: 1876  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-12745

Query Match 57.9%; Score 1349.4; DB 5; Length 1876;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 1749; Conservative 0; Mismatches 21; Indels 367; Gaps 2;  
QY 196 CAAACAGTATCTTCACCGAAGTGAACCCAGGCGCAAAATTTGAGTCCCTTTTAGGA 255  
DB 60 CAATCTGCATTTTGAAGAAAGCAATTTTGGGCCAAATTTGAGTCCCTTTTAGGA 119  
QY 256 CGTATGACAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAAAAGGAGTCAGATTAAC 315  
DB 120 CGTATGACAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAAAAGGAGTCAGATTAAC 179  
QY 316 TCAGCAATCCCTTCTCCGAGCAGATGCCAGCTCCAGCTCCAGCTGCATTAAGACTGAGTTTCTGG 375  
DB 180 TCAGCAATCCCTTCTCCGAGCAGATGCCAGCTCCAGCTCCAGCTGCATTAAGACTGAGTTTCTGG 239  
QY 376 GAAAGCAATGAAGTTATATTTTCTCAGACTTACATATGATCTCTTATATGCCATCTCCAAGC 435  
DB 240 GAAAGCAATGAAGTTATATTTTCTCAGACTTACATATGATCTCTTATATGCCATCTCCAAGC 299  
QY 436 CGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCTCCGCTCTCCGCTCTCCGCTAGTGGATGGA 495  
DB 300 CGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCTCCGCTCTCCGCTAGTGGATGGA 359  
QY 496 AACAGTGGAGATGCCAGCTTCAATTAATCTATGATCTCTTATATGCCATCTCCAAGC 555  
DB 360 AACAGTGGAGATGCCAGCTTCAATTAATCTATGATCTCTTATATGCCATCTCCAAGC 419  
QY 556 TGGGCGCAGGGAAGATGATGAAATTCAGCAGGAGCTGACCCACTTCCAGCTCCAGCTGGTG 615  
DB 420 TGGGCGCAGGGAAGATGATGAAATTCAGCAGGAGCTGACCCACTTCCAGCTCCAGCTGGTG 479  
QY 616 TCCATGATGTGAGTGTATCAAGAGAGGAGGAAGAGGAAATGGAAGATGAGGA 675  
DB 480 TCCATGATGTGAGTGTATCAAGAGAGGAGGAAGAGGAAATGGAAGATGAGGA 539  
QY 676 GACCTAAGCCAAATATCCAGCAGGAGGCGGAGTACAGCCGATCCACTCAGCT 735  
DB 540 GACCTAAGCCAAATATCCAGCAGGAGGCGGAGTACAGCCGATCCACTCAGCT 599  
QY 736 GAACTGCACGCGACGAGGACGCAATCCAAATCATCTACGCGGAGGAATCTTTTACTGT 795  
DB 600 GAACTGCACGCGACGAGGACGCAATCCAAATCATCTACGCGGAGGA----- 648  
QY 796 GGAGTGGCTGTACGACTTTCTTCGAGGTGGCAGCCGAGATCGGGGTGGCAGAAATCC 855  
DB 649 ----- 648  
QY 856 CAGTTTATGTTGCTCAGAGAGAAATCAAGCCGCTGTCCTTCTAATGTCTGCACACC 915  
DB 649 ----- 648  
QY 916 AGTTACTGTTTCAGCCCGGGAATGACTTGGGCCAATCACTGAGTTTGGTGTATCCG 975  
DB 649 ----- 648  
QY 976 ACAAGGACATTTGGGACTGTCTTGAGAAACACAGATAATGATGTGTTTGTACTTTTCT 1035  
DB 649 ----- 648  
QY 1036 TTTCTGTAGTTTCTGTCTGTGCCAAGGCGAGTTGATCAGTGAGCTCAGGAGAGCTT 1095  
DB 649 ----- 648  
QY 1096 CTTGTTTCTAAGTGGCTCGAGGGGCCACTCTTACTGTAGGAAGAGGTACACAGGA 1155

DB 649 ----- 648  
QY 1156 GCGCCCTAGTGACAGAGGTTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTACCGTT 1215  
DB 654 GCGCCCTAGTGACAGAGGTTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTACCGTT 713  
QY 1216 TCTTTTCTCCCTCATGTTTCTCATGTTTGTGCAATGTATATTTTACAGACTAAC 1275  
DB 714 TCTTTTCTCCCTCATGTTTCTCATGTTTGTGCAATGTATATTTTACAGACTAAC 773  
QY 1276 CTTTGTTCGTATATAAGTATACACCGTTTGTGTTTACATCTTTTGGGAAGCAGGAAG 1335  
DB 774 CTTTGTTCGTATATAAGTATACACCGTTTGTGTTTACATCTTTTGGGAAGCAGGAAG 833  
QY 1336 CGTTTGGAAAACGTATCACCTTTTCCAGATTTCTCGATTTCTCGACTCTTTGCAACAGCAC 1395  
DB 834 CGTTTGGAAAACGTATCACCTTTTCCAGATTTCTCGATTTCTCGACTCTTTGCAACAGCAC 893  
QY 1396 TTGCTTTGGGAACTCTTCTTGAATGCAATTTCACTCAGCATCCCAACCGTGCAACGTGTA 1455  
DB 894 TTGCTTTGGGAACTCTTCTTGAATGCAATTTCACTCAGCATCCCAACCGTGCAACGTGTA 953  
QY 1456 ACTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATTTCTCTGTAGAAATTTAGCTTATACAA 1515  
DB 954 ACTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATTTCTCTGTAGAAATTTAGCTTATACAA 1013  
QY 1516 TTCAGAGAAATAGCAGTTTCACTGCCAACTTTTAGTGGGTGAGAAATTTTAGTTAGTGT 1575  
DB 1014 TTCAGAGAAATAGCAGTTTCACTGCCAACTTTTAGTGGGTGAGAAATTTTAGTTAGTGT 1073  
QY 1576 TTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTATGTTGGTGGTTTCTATACATGAATCA 1635  
DB 1074 TTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTATGTTGGTGGTTTCTATACATGAATCA 1133  
QY 1636 TAGCCAAAACCTTTTGG-AAAAGTGTGTTGAGATAGTTGTTTCTTTTACCCCAAGAA 1694  
DB 1134 TAGCCAAAACCTTTTGGAAAACCTGTTGTTGAGATAGTTGTTTCTTTTACCCCAAGAA 1193  
QY 1695 GACATCAAGATACACTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1754  
DB 1194 GACATCAAGATACACTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1253  
QY 1755 GGGTGAAGATATGCGAGTGGGAGCTAAGATGATTAATTAACCTACCTGTGAATCATGTT 1814  
DB 1254 GGGTGAAGATATGCGAGTGGGAGCTAAGATGATTAATTAACCTACCTGTGAATCATGTT 1313  
QY 1815 GTAGAAAAGCTGTTCCCATGTCTTAACAGGACTTGAATTCAGAGCATGTCAAGTGGATAG 1874  
DB 1314 GTAGAAAAGCTGTTCCCATGTCTTAACAGGACTTGAATTCAGAGCATGTCAAGTGGATAG 1373  
QY 1875 TAGATCTGTGGCGATATGAGAGGATGCGAGTCCCTTTCCCATTCATCTCTGATGGAAT 1934  
DB 1374 TAGATCTGTGGCGATATGAGAGGATGCGAGTCCCTTTCCCATTCATCTCTGATGGAAT 1433  
QY 1935 GTTATCTAGTTTAAACATTTGTAATTTTCTAGTTGTAATGTTGTAATGTTGTAATGTTGTAAT 1994  
DB 1434 GTTATCTAGTTTAAACATTTGTAATTTTCTAGTTGTAATGTTGTAATGTTGTAATGTTGTAAT 1493  
QY 1995 GGTATTTATTTTGGCCCTTACAATACCGTAACCAATGTTTGTCAATTTTGAATACTTAAATG 2054  
DB 1494 GGTATTTATTTTGGCCCTTACAATACCGTAACCAATGTTTGTCAATTTTGAATACTTAAATG 1553  
QY 2055 CCAAGTAAACATTCATGCTTTGGAATTTTGAAGATGTTTATTTCTTTGAGAGCAAT 2114  
DB 1554 CCAAGTAAACATTCATGCTTTGGAATTTTGAAGATGTTTATTTATTTTGAAGCAAT 1613  
QY 2115 ATGTTTGCATTAATGCTTTGATTGTTTCAATCAAGAAATTTGATTTGAAGCTTCTCAACC 2174  
DB 1614 ATGTTTGCATTAATGCTTTGATTGTTTCAATCAAGAAATTTGATTTGAAGCTTCTCAACC 1673  
QY 2175 CTGTTTACGCTACTTGGTAAGAGGAGCGGTTTGGGAGAGCAATTTGATCGTGTCCA 2234

```

Db 1674 CTGTTTACGTAAGTGTGAAGAGGAGCGGTTTGGAGAGACCATTCATCGCTGTCCA 1733
Qy 2235 AGTGTTCCTTGTAAAGTCTTTTAACTCGAGAGGCTTAACCTCAAAAATACCTTTTAAAC 2294
Db 1734 AGTGTTCCTTGTAAAGTGTGTAAAGTGAAGAGGCTTAACCTCAAAAATATTTTAAAC 1793
Qy 2295 TGCATTCCTATATAATAGGAGCAGATGCTCTTAC 2331
Db 1794 TGCATTCCTATATAATAGGAGCAGATGCTCTTAC 1830

RESULT 13
US-09-724-676A-12745
; Sequence 12745, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patemlin version 3.2
; SEQ ID NO 12745
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12745

Query Match 57.9%; Score 1349.4; DB 5; Length 1876;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 21; Indels 367; Gaps 2;

Qy 196 CAACAGGTATCTTACGAGAAAGTGAACAGGCGCAAAATTTGAGTCCCTCTTTAGGA 255
Db 60 CAATCTGACATATTTGAAGAAAGGAAAGGCAATGTTGGGCAAAATTTGAGTCCCTCTTTAGGA 119
Qy 256 CGTATGCAAGAGCATCACTCTTCAATTTTAAAGCTTCAAAAGAGTCAAAATTAAT 315
Db 120 CGTATGCAAGAGCATCACTCTTCAATTTTAAAGCTTCAAAAGAGTCAAAATTAAT 179
Qy 316 TCAGCAACCCCTTCTCCGAGAGAGTCCGAGGCTCCAGCTGCAATGAAGTCTTCTGG 375
Db 180 TCAGCAACCCCTTCTCCGAGAGAGTCCGAGGCTCCAGCTGCAATGAAGTCTTCTGG 299
Qy 376 GAAAGAAATGAAGTATATTTTGTCTAGACCTTACATAGAAAGTCAACCTGCTC 435
Db 240 GAAAGAAATGAAGTATATTTTGTCTAGACCTTACATAGAAAGTCAACCTGCTC 299
Qy 436 CGCCAAATCCAGACAGAGCTTTCTGATCTTCCCTCCGCTCCGCGCCAGTGGAGTGA 495
Db 300 CGCCAAATCCAGACAGAGCTTTCTGATCTTCCCTCCGCTCCGCGCCAGTGGAGTGA 389
Qy 496 AACCAAGTGAAGATGCGACCCCGCATTAACATATGATCTCTTATATGCAATCCAAAGC 555
Db 360 AACCAAGTGAAGATGCGACCCCGCATTAACATATGATCTCTTATATGCAATCCAAAGC 419
Qy 556 TGGGGCCAGGGGAAAGATATGCAACGAGGCACTGACCACTCCAGCGTGTGG 615
Db 420 TGGGGCCAGGGGAAAGATATGCAACGAGGCACTGACCACTCCAGCGTGTGG 479
Qy 616 TCCATGTATGTGAGAGTATCAAGAGAGAGAGAAAGAAAGAAATGGAATGAGGA 675
Db 480 TCCATGTATGTGAGAGTATCAAGAGAGAGAGAGAAAGAAAGAAATGGAATGAGGA 539
Qy 676 GACCTTAAGCAAAAATTTATCCAGACAGAGGCGGAGTACACGCGATCCAGCTCAAGCT 735
Db 540 GACCTTAAGCAAAAATTTATCCAGACAGAGGCGGAGTACACGCGATCCAGCTCAAGCT 599
Qy 736 GAACTGCGACGCGAGAGAGCGCATTTCCAAATCATATCTCAAGGAGAAATCTTTACTGT 795
Db 600 GAACTGCGACGCGAGAGAGCGCATTTCCAAATCATATCTCAAGGAGAAATCTTTACTGT 648
Qy 796 GGAGGTGGCTGATCAAGACTTCTTGAGAGTGGCAAGCCGAGATCGGGGTGGCAGAAATCC 855

```

```

Db 649 ----- 648
Qy 856 CAGTTCATGTGCTCAGAGAGAAATCAAGCGGTGTCCCTGTGTTCAATGCTGCACACC 915
Db 649 ----- 648
Qy 916 AGTTACTGTTCAATGACCCCGGAATGACTTGGGCCAATCAGTGAATTTGTGTATGCC 975
Db 649 ----- 648
Qy 976 ACAAGACATTGGGACTGTCTTGAGAAAAAGATATATAGATGTGTTGTACTTGTCT 1035
Db 649 ----- 648
Qy 1036 TTTCTGTAGGTTCTGTGTGTGCCAAGGCGAGGTGATCATGAGCTCAGAGAGAGCTT 1095
Db 649 ----- 648
Qy 1096 CTTGTTCTAAGTGGCCTGAGGGGCACTCTACTGTGAGGAGAGTACCAAGAA 1155
Db 649 -----AGGA 653
Qy 1156 GCCGCTAGTGCAGAGAGGTTTGAAAAACAGACCAATGCAATGTGAAAATTGAGCGTT 1215
Db 654 GCCGCTAGTGCAGAGAGGTTTGAAAAACAGACCAATGCAATGTGAAAATTGAGCGTT 713
Qy 1216 TCCCTTCTCCCTCATGTTCTCATGTTTGTGATGTATATTACTGATTTCAAGACTAAC 1275
Db 714 TCCCTTCTCCCTCATGTTCTCATGTTTGTGATGTATATTACTGATTTCAAGACTAAC 773
Qy 1276 CTTTGTTGTTATTAAGTTACACCGTTGTTGTTTACATCTTTTGGGAAGCCAGGAAG 1335
Db 774 CTTTGTTGTTATTAAGTTACACCGTTGTTTACATCTTTTGGGAAGCCAGGAAG 833
Qy 1336 CGTTTGGAAAAAGTATCACTTTCCAGATTTCGGAATTCGACTCTTTGGCAACAGAC 1395
Db 834 CGTTTGGAAAAAGTATCACTTTCCAGATTTCGGAATTCGACTCTTTGGCAACAGAC 893
Qy 1396 TTGCTTGGGAACTTTCTTCGGAATTCGACTCTTCAGATCCCAACCGGCAAGTGA 1455
Db 894 TTGCTTGGGAACTTTCTTCGGAATTCGACTCTTCAGATCCCAACCGGCAAGTGA 953
Qy 1456 ACTTGTCTTTTGGAAAAAGTATGATGAAATTCCTCTGTAGAAATTTAGCTTATACAA 1515
Db 954 ACTTGTCTTTTGGAAAAAGTATGATGAAATTCCTCTGTAGAAATTTAGCTTATACAA 1013
Qy 1516 TTCAAGAAATAGCACTTTCAGTCCCACTTTTATGTGGTGAATAATTTAGTTAGTGT 1575
Db 1014 TTCAAGAAATAGCACTTTCAGTCCCACTTTTATGTGGTGAATAATTTAGTTAGTGT 1073
Qy 1576 TTGGGATGGAGCTGAGTTTCTGTGTTCTTTATGTGGTGTTCATACATGATCA 1635
Db 1074 TTGGGATGGAGCTGAGTTTCTGTGTTCTTTATGTGGTGTTCATACATGATCA 1133
Qy 1636 TAGCCAAAACTTTTGTGG-AAACTGTGTTGAGATGTTGTTCTTTTATCCACGAA 1694
Db 1134 TAGCCAAAACTTTTGTGGAAACTGTGTTGAGATGTTGTTCTTTTATCCACGAA 1193
Qy 1695 GACATCAAGATACCTTGTAAATTAAGCTGATAGATATATCAATCTGTGTACACTT 1754
Db 1194 GACATCAAGATACCTTGTAAATTAAGCTGATAGATATATCAATCTGTGTACACTT 1253
Qy 1755 GGGTGAAGATATGAGAGGAGAGCTAAGATGTTAATCACTACCTGTGAATCATATGTT 1814
Db 1254 GGGTGAAGATATGAGAGGAGAGCTAAGATGTTAATCACTACCTGTGAATCATATGTT 1313
Qy 1815 GTAGAAAAAGCTGTTCCATGTCTTAAAGAGACTTGAATTTCAAGATGTCAAGTGTATG 1874
Db 1314 GTAGAAAAAGCTGTTCCATGTCTTAAAGAGACTTGAATTTCAAGATGTCAAGTGTATG 1373
Qy 1875 TAGATCTGTGGCATATAGAGGATGCAAGTGTCTTTCCCAATGATTCATGATGAAATT 1934

```

Db 1374 TAGATCTCTGGCGATATGAGAGGATCGAGTGCCTTTCCCAATTCCTGATGAAAT 1433  
Qy 1935 GTTATACCTAGGTTAACTTTTCTAGTTGTAATGTAATGTAATGTAATGTAATGTAAT 1994  
Db 1434 GTTATACCTAGGTTAACTTTTCTAGTTGTAATGTAATGTAATGTAATGTAATGTAAT 1493  
Qy 1995 GGTATATATTTGGCCCTTCAATACCGTAACAAATGTTGTCATTTTGAATACTTTAATG 2054  
Db 1494 GGTATATATTTGGCCCTTCAATACCGTAACAAATGTTGTCATTTTGAATACTTTAATG 1553  
Qy 2055 CCAAGTAACAAATGATCTTTTGAATAATTTGGAAGATGTTTATCTTTGAGAACCAAT 2114  
Db 1554 CCAAGTAACAAATGATCTTTTGAATAATTTGGAAGATGTTTATCTTTGAGAACCAAT 1613  
Qy 2115 ATGTTTTCATTAATGCTTTGATGTTTCAATCAAGAAATGATGAAAGCTTCTCAAAACC 2174  
Db 1614 ATGTTTTCATTAATGCTTTGATGTTTCAATCAAGAAATGATGAAAGCTTCTCAAAACC 1673  
Qy 2175 CTGTTTACGCTACTTGGTAAGAGGAGCGGTTTGGGAGAGCACTTGCATCGCTGTCCA 2234  
Db 1674 CTGTTTACGCTACTTGGTAAGAGGAGCGGTTTGGGAGAGCACTTGCATCGCTGTCCA 1733  
Qy 2235 AGTTTCTCTGTAAGTCTTTTAACTGGAGAGGCTAACCTCAAAATACCTTTTAAAC 2294  
Db 1734 AGTTTCTCTGTAAGTCTTTTAACTGGAGAGGCTAACCTCAAAATACCTTTTAAAC 1793  
Qy 2295 TGCATTTCTATAATAATGGGACAGTATGCTCTTAC 2331  
Db 1794 TGCATTTCTATAATAATGGGACAGTATGCTCTTAC 1830

## RESULT 14

US-09-724-676-12743

; Sequence 12743, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 12743

; LENGTH: 1839

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-724-676-12743

Query Match 57.8%; Score 1348.4; DB 5; Length 1839;  
Best Local Similarity 82.5%; Pred. No. 0;  
Matches 1736; Conservative 0; Mismatches 1; Indels 367; Gaps 2;

Qy 229 GGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTTCAGTATTTTA 288  
Db 56 GGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTTCAGTATTTTA 115  
Qy 289 AGAGCTTCAAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGACAGATGCCAGGC 348  
Db 116 AGAGCTTCAAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGACAGATGCCAGGC 175  
Qy 349 TCCAGCTGCATAGACTGAGTTCTCGGAAAGGAATGAAATGAAATGAAATGAAATGAAATGAAAT 408  
Db 176 TCCAGCTGCATAGACTGAGTTCTCGGAAAGGAATGAAATGAAATGAAATGAAATGAAATGAAAT 235  
Qy 409 TACACATAGGAAGCTCAGACTGGCTCCGCAAAATCCAGACAGGAGTTTCTGATCTCTCC 468  
Db 236 TACACATAGGAAGCTCAGACTGGCTCCGCAAAATCCAGACAGGAGTTTCTGATCTCTCC 295  
Qy 469 CTCCCGCTCTCCGCGAGTGGGATGGAACAGTGGGAAGTGGGATGCGACCCAGCTATAACT 528  
Db 296 CTCCCGCTCTCCGCGAGTGGGATGGAACAGTGGGAAGTGGGATGCGACCCAGCTATAACT 355  
Qy 529 ATGATCTCTTATATGCCATCTCCAGCTGGGCGCAGGGAAGTATGAAATGCAACGAG 588

Db 356 ATGATCTCTTATATGCCATCTCCAAGCTGGGCGAGGGAAGTATGAATTCACGCGAG 415  
Qy 589 CGACTGACACCACTCCAGCGTGGTGGTCCATGATGTGAGAGTATCAAGAGAAGGAGG 648  
Db 416 CGACTGACACCACTCCAGCGTGGTGGTCCATGATGTGAGAGTATCAAGAGAAGGAGG 475  
Qy 649 AAGAAGAGGAATGGAAGAATGAGGAGACTTAAGCCAAATAATTCAGACACGAGGCG 708  
Db 476 AAGAAGAGGAATGGAAGAATGAGGAGACTTAAGCCAAATAATTCAGACACGAGGCG 535  
Qy 709 CGAGGTACACCGCGATCCACCTCAGTGAACCGGACGAGGAGGACGATTCCTCAAAATC 768  
Db 536 CGAGGTACACCGCGATCCACCTCAGTGAACCGGACGAGGAGGACGATTCCTCAAAATC 595  
Qy 769 ATACTCAGCGGAGGAATCTTTTACTGTGAGGTGGTGGTCCAGACTTCTCGAGGTTGG 828  
Db 596 ATACTCAGCGGAGGA----- 611  
Qy 829 CAGCCGAGATCGGGGTGGGAGAAATCCCAAGTTTCTCAGAAAGAGATCAAGGCGG 888  
Db 612 ----- 611  
Qy 889 TGTCCCTTTGTTCTAATGCTGCACACAGTACTTCTCATGGCACCCGGGAATGATCTGG 948  
Db 612 ----- 611  
Qy 949 GCAATCACTGAGTTTGTGGTGAATGCAAGGACATTTGGGACTGTCTTGAGAAACAG 1008  
Db 612 ----- 611  
Qy 1009 ATAATGATAGTGTTTGTACTTGTCTTTCTGGTAGGTTCTGTCTGTGCCAAGGCGAGG 1068  
Db 612 ----- 611  
Qy 1069 TTGATCAGTGAAGCTCAGGAGAGAGCTTCTGTCTTAAGTGGCCCTGCAGGGGCCACTCTC 1128  
Db 612 ----- 611  
Qy 1129 TACTGTAGGAAGAGGTACCAAGAGGCGCTTAGTGCAGAGAGGTTGTGAAAAACAGCA 1188  
Db 612 -----AGGAAGCGCCCTAGTGCAGAGAGGTTGTGAAAAACAGCA 649  
Qy 1189 GCAATGCAATGTGGAATTTGTAGCGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCA 1248  
Db 650 GCAATGCAATGTGGAATTTGTAGCGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCA 709  
Qy 1249 TGTATATTAATGATTTTCAAGACTTAACCTTTTGTCTGTATATAAGTTTACACCGTCTGTCT 1308  
Db 710 TGTATATTAATGATTTTCAAGACTTAACCTTTTGTCTGTATATAAGTTTACACCGTCTGTCT 769  
Qy 1309 TTTACATCTTTTGGGAGCCAGGAAGCGTTTGGAAAAAGTATCACCTTTTCCAGATTCT 1368  
Db 770 TTTACATCTTTTGGGAGCCAGGAAGCGTTTGGAAAAAGTATCACCTTTTCCAGATTCT 829  
Qy 1369 CGGATCTCGACTCTTTGCAACAGCACTTGTCTGCGAAAGTATGCTTTTGGAAAAAGTATGATCTGAAA 1428  
Db 830 CGGATCTCGACTCTTTGCAACAGCACTTGTCTGCGAAAGTATGCTTTTGGAAAAAGTATGATCTGAAA 889  
Qy 1429 TCAGCATCCCAACCGTGCAGAGTAACTTGTGCTTTTGGAAAAAGTATGATCTGAAA 1488  
Db 890 TCAGCATCCCAACCGTGCAGAGTAACTTGTGCTTTTGGAAAAAGTATGATCTGAAA 949  
Qy 1489 TTCCTCTGTAGAAATTTAGCTTATCAAAATTCAGAGAAATAGCAGTTTCTACGCAACTTTTA 1548  
Db 950 TTCCTCTGTAGAAATTTAGCTTATCAAAATTCAGAGAAATAGCAGTTTCTACGCAACTTTTA 1009  
Qy 1549 GTGGGTGAGAAATTTTGTAGTTTGGGATCGGACCTCAGTTTCTGTTTCTTTT 1608  
Db 1010 GTGGGTGAGAAATTTTGTAGTTTGGGATCGGACCTCAGTTTCTGTTTCTTTT 1069  
Qy 1609 TATGTGTGTGTTTCTATACATGAAATCATAGCAAAATCTTTTGG-AAAATGTTGGTTG 1667



Db 770 TTATCATCTTTTGGAGCCAGGAAGCGTTTGGAAACGTATACCTTTCCAGATTCT 829  
Qy 1369 CGGATTCCTGACTCTTTTGCACAGCACTGCTTGGGAACTCTTCTGGAAATGCATTAC 1428  
Db 830 CGGATTCCTGACTCTTTTGCACAGCACTGCTTGGGAACTCTTCTGGAAATGCATTAC 889  
Qy 1429 TCAGCATCCCAACCGTGCACCGTGAACCTTGTGCTTTTGCACAAAGAGTTGATCTGAAA 1488  
Db 890 TCAGCATCCCAACCGTGCACCGTGAACCTTGTGCTTTTGCACAAAGAGTTGATCTGAAA 949  
Qy 1489 TTCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACGCCAACTTTTA 1548  
Db 950 TTCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACGCCAACTTTTA 1009  
Qy 1549 GTGGGTGAGAAATTTTGTAGTTTGTAGTGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTT 1608  
Db 1010 GTGGGTGAGAAATTTTGTAGTTTGTAGTGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTT 1069  
Qy 1609 TATGTGGTGGTCTTCTATACATGAATCATAGCCAAACCTTTTGGG-AAAAGCTGTTG 1667  
Db 1070 TATGTGGTGGTCTTCTATACATGAATCATAGCCAAACCTTTTGGG-AAAAGCTGTTG 1129  
Qy 1668 AGATAGTGGTCTTCTTACCCACGAGACATCAAGATACACTTGTAAATAAGCTGATA 1727  
Db 1130 AGATAGTGGTCTTCTTACCCACGAGACATCAAGATACACTTGTAAATAAGCTGATA 1189  
Qy 1728 GCATATATTATACACCTGTTGTACACTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATG 1787  
Db 1190 GCATATATTATACACCTGTTGTACACTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATG 1249  
Qy 1788 TATTAACTTACCTGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGCTCTAACAGGACT 1847  
Db 1250 TATTAACTTACCTGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGCTCTAACAGGACT 1309  
Qy 1848 TCAATTTCAAAGCATGCTCAAGTGGATAGATCTGTGGCGATATGAGAGGATGCAGTGC 1907  
Db 1310 TCAATTTCAAAGCATGCTCAAGTGGATAGATCTGTGGCGATATGAGAGGATGCAGTGC 1369  
Qy 1908 CTTTCCCATTCATTCTCTGATGGAATTTGTATAGTTTAAACATTTGTAATTTTTTCT 1967  
Db 1370 CTTTCCCATTCATTCTCTGATGGAATTTGTATAGTTTAAACATTTGTAATTTTTTCT 1429  
Qy 1968 AGTTGTAATGTATGCTGTGTAATAGTATATATTTTGGCCTTACAAATACCGTAACA 2027  
Db 1430 AGTTGTAATGTATGCTGTGTAATAGTATATATTTTGGCCTTACAAATACCGTAACA 1489  
Qy 2028 ATGTTTGTCAITTTGAAATACCTTAATGCCAAGTAAAGTGCATGCTTTGGAATTTGGAA 2087  
Db 1490 ATGTTTGTCAITTTGAAATACCTTAATGCCAAGTAAAGTGCATGCTTTGGAATTTGGAA 1549  
Qy 2088 GATGGTTTATTTCTTTGAGAAGCAAAATATGTTTGCATTTAAATGCTTTGATTTTCATATC 2147  
Db 1550 GATGGTTTATTTCTTTGAGAAGCAAAATATGTTTGCATTTAAATGCTTTGATTTTCATATC 1609  
Qy 2148 AAGAAATTTGATTTGAACTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGGAGCGGTT 2207  
Db 1610 AAGAAATTTGATTTGAACTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGGAGCGGTT 1669  
Qy 2208 TGGGAGAGACCATTTGCATCGCTGCCAAGTGTCTTGTAAAGTCTTTTAACTGGAGA 2267  
Db 1670 TGGGAGAGACCATTTGCATCGCTGCCAAGTGTCTTGTAAAGTCTTTTAACTGGAGA 1729  
Qy 2268 GGTAAACCTCAAAATACTTTTAACTGATTTTAACTGATTTTAACTGATTTTAACTGATTT 2327  
Db 1730 GGTAAACCTCAAAATACTTTTAACTGATTTTAACTGATTTTAACTGATTTTAACTGATTT 1789  
Qy 2328 TTAC 2331  
Db 1790 TTAC 1793

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 29.8988 Seconds  
(without alignments)  
2020.659 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACVVDV.....SPKPKIQTTRRGLPPSVSN 197

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DB=us09-782-953-9  
-O=/cg2\_1/USPTO.spool/US09782953/runat\_11122002\_114430\_17377/app\_query.fasta\_1.1173  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 @CGN 1 1 29 @runat\_11122002\_114430\_17377 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOP -DBV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:  
1: /cg2\_6/ptodata/1/ina/5A COMB seq.\*  
2: /cg2\_6/ptodata/1/ina/5B COMB seq.\*  
3: /cg2\_6/ptodata/1/ina/6A COMB seq.\*  
4: /cg2\_6/ptodata/1/ina/6B COMB seq.\*  
5: /cg2\_6/ptodata/1/ina/PTUS COMB seq.\*  
6: /cg2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description	
	Score	Match	Length				
1	600	57.6	2174	2	US-08-665-040-1	Sequence 1, Appli	
2	101	9.7	5115	3	US-08-348-518C-3	Sequence 3, Appli	
3	101	9.7	5115	3	US-08-476-509B-3	Sequence 3, Appli	
c	4	89.5	8.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	5	89	8.5	3211	2	US-08-574-959A-8	Sequence 8, Appli
	6	89	8.5	3211	4	US-09-357-014-8	Sequence 8, Appli
	7	89	8.5	3901	2	US-08-574-959A-6	Sequence 6, Appli
8	89	8.5	3901	4	US-09-357-014-6	Sequence 6, Appli	
9	83.5	8.0	1869	3	US-09-372-668-3	Sequence 3, Appli	
c	10	82.5	7.9	3483	3	US-08-808-599A-23	Sequence 23, Appli
11	82.5	7.9	4411529	4	US-09-103-840A-1	Sequence 1, Appli	
12	82	7.9	4112	1	US-08-340-203A-2	Sequence 2, Appli	

13	82	7.9	4112	2	US-08-452-567-2	Sequence 2, Appli
14	82	7.9	4112	2	US-08-452-427-2	Sequence 2, Appli
15	82	7.9	4112	3	US-09-085-407-2	Sequence 2, Appli
16	82	7.9	4616	1	US-08-340-203A-1	Sequence 1, Appli
17	82	7.9	4616	2	US-08-452-567-1	Sequence 1, Appli
18	82	7.9	4616	2	US-08-452-427-1	Sequence 1, Appli
19	82	7.9	4616	3	US-09-085-407-1	Sequence 1, Appli
20	80	7.7	3089	1	US-07-921-796-7	Sequence 7, Appli
21	79.5	7.6	5410	4	US-09-221-017B-70	Sequence 70, Appli
22	79	7.6	169998	4	US-09-676-610B-24	Sequence 24, Appli
23	78	7.5	3360	1	US-08-408-093-5	Sequence 5, Appli
24	78	7.5	3360	1	US-08-408-420A-5	Sequence 5, Appli
25	78	7.5	3360	1	US-08-714-901-5	Sequence 5, Appli
26	78	7.5	3360	3	US-08-040-741-5	Sequence 5, Appli
27	78	7.5	3468	1	US-07-951-715A-2	Sequence 2, Appli
28	78	7.5	3468	1	US-07-951-715A-4	Sequence 4, Appli
29	78	7.5	3468	2	US-08-459-448A-2	Sequence 2, Appli
30	78	7.5	3468	2	US-08-459-448A-4	Sequence 4, Appli
31	78	7.5	3468	3	US-08-459-595A-2	Sequence 2, Appli
32	78	7.5	3468	3	US-08-459-595A-4	Sequence 4, Appli
33	78	7.5	3468	3	US-08-459-504B-2	Sequence 2, Appli
34	78	7.5	3468	3	US-08-459-504B-4	Sequence 4, Appli
35	78	7.5	3468	3	US-08-459-444-2	Sequence 2, Appli
36	78	7.5	3468	3	US-08-459-444-4	Sequence 4, Appli
37	78	7.5	3468	3	US-09-053-549-3	Sequence 3, Appli
38	78	7.5	3468	3	US-09-053-549-5	Sequence 5, Appli
39	78	7.5	3468	4	US-09-547-422-2	Sequence 2, Appli
40	78	7.5	3468	4	US-09-547-422-4	Sequence 4, Appli
C 41	78	7.5	80161	3	US-09-036-987A-1	Sequence 1, Appli
C 42	78	7.5	80161	4	US-09-370-700-1	Sequence 1, Appli
C 43	77.5	7.4	2524	1	US-08-317-522A-1	Sequence 1, Appli
C 44	77.5	7.4	2524	1	US-08-439-818A-1	Sequence 1, Appli
C 45	77.5	7.4	2524	2	US-08-751-965-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-665-040-1  
; Sequence 1, Application US/08665040  
; Patent No. 5869318  
; GENERAL INFORMATION:

APPLICANT: ESTIVILL PALLEJA, XAVIER  
APPLICANT: FUENTES, JUAN JOSE  
APPLICANT: PRITCHARD, MELANIE  
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,  
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,  
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PARRY  
STREET: 26 WEST 61ST STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10023

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1 FOR DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,040  
FILING DATE: JUNE 7, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: ES P9501140  
FILING DATE: JUNE 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: JANET I. CORD

1	REGISTRATION NUMBER: 33,778	
2	REFERENCE/DOCKET NUMBER: U010815-9	
3	TELECOMMUNICATION INFORMATION:	
4	TELEPHONE: (212) 708-1800	
5	TELEFAX: (212) 246-8959	
6	INFORMATION FOR SEQ ID NO: 1:	
7	SEQUENCE CHARACTERISTICS:	
8	LENGTH: 2174 base pairs	
9	TYPE: nucleic acid	
10	STRANDEDNESS: double	
11	TOPOLOGY: circular	
12	MOLECULE TYPE: cDNA for mRNA	
13	HYPOTHETICAL: No	
14	ANTI-SENSE: No	
15	ORIGINAL SOURCE: human	
16	ORGANISM: Homo sapiens	
17	DEVELOPMENTAL STAGE: foetal	
18	TISSUE TYPE: Brain	
19	IMMEDIATE SOURCE: gene library of cDNA	
20	LIBRARY: gene library of cDNA from foetal	
21	CLONE: BC-17.8-1 and BC-17.8-2	
22	POSITION IN GENOME:	
23	CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9	
24	MAP POSITION: 21q22.1-q22.2	
25	FEATURE:	
26	NAME/KEY: cDNA for mRNA, BC-17.8	
27	LOCATION: 1..2174	
28	FEATURE:	
29	NAME/KEY: untranslated 5'	
30	LOCATION: 1..48	
31	FEATURE:	
32	NAME/KEY: coding sequence	
33	LOCATION: 49..560	
34	OTHER INFORMATION: Down Syndrome critical	
35	FEATURE:	
36	NAME/KEY: DSCR1	
37	LOCATION: 1..171 PEPTIDES	
38	IDENTIFICATION METHOD: translation of the	
39	OTHER INFORMATION:	
40	OTHER INFORMATION: - deduced protein	
41	OTHER INFORMATION: - proline-rich protein domains	
42	OTHER INFORMATION: - glutamic acid-rich protein domains	
43	OTHER INFORMATION: - leucine/phenylalanine-rich protein domains	
44	FEATURE:	
45	NAME/KEY: untranslated 3'	
46	LOCATION: 564..2174	
47	FEATURE:	
48	NAME/KEY: 2 poly (A)	
49	LOCATION: 1541..1546 AND 2132..2137	
50	US-08-665-040-1	
51	Alignment Scores:	
52	Pred. No.: 1.74e-58	
53	Score: 600.00	
54	Percent Similarity: 79.77%	
55	Best Local Similarity: 66.47%	
56	Query Match: 57.64%	
57	DB: 2	
58	Gaps: 2	
59	US-09-782-953-9 (1-197) x US-08-665-040-1 (1-2174)	
60	OY	
61	25 AengingluVallysglyluePheglu1yeuPheargThryAePgluCyAValThr 44	
62	43 AACAGAAATGGTGTATAGCCAAATTTAGTCCTCTTTAGACGATATGACACAGACATCC 102	
63	OY	
64	45 Pheglu1yeuPheylserPheargArgValArgileanPheSerHieProlysSerila 64	
65	103 TTTGCTATTTTAAAGGCTTCAMAGAGTCGAAATMAACTTCACCAACCCCTTCGCCA 162	
66	OY	
67	65 A1aAga1AaRg11eGlu1euh1eGlu1hRg1nPhaRg1yLyV1eLyLeu1eLyTyr 84	
68	163 GCAGATGCCAGGCTCCAGCTGCA1TAAGACT1GAGTTTCGGGAAAGAAATTAAGTTAT 222	

```

0y 85 PheaaGlnVnIaGlnThrProGluThrAspGlyAspLeuHisLeuAlaProProGln 104
Db 223 TTTCGTCAAGACCTTACACATACAGAAAGCTCA-----CACCTGGCTCCGCCAAAT 270
0y 105 ProAlaValGlnPheLeuLeuIleSerProProSerSerProProValGlyTyrPlysProIle 124
Db 271 CCAGCAAGACGATTTCTTGATCTCTCCCTCCGCCCTCTCCGACAGTGGGATGGAAACAACGTG 330
0y 125 SerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaValGlyPro 144
Db 331 GAAGATGGACCCCAAGCATATMACTATGATCTCTTATATGCACTCTCAAGCTGGGGCCA 390
0y 145 GlyGluTyrTyrGluLeuHisAlaGlyThrGluSerThrProSerValValValHisVal 164
Db 391 GGGGAAAGATGATGATCAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGACCTTACGTA 450
0y 165 CysAspSerAspMetGluGluGluGluGluAspProLysThrSer-----Prolys 180
Db 451 TGTGAGACTGATCAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGACCTTACGTA 510
0y 181 ProlysIleIleGlnThrThrArgProGlyLeuProPro 193
Db 511 CCAAAATTATTCAGACCAAGCGCCGGAGTACACCGCG 549

RESULT 2
US-08-348-518C-3
; Sequence 3, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Klauder & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:

```

CLONE: VAP  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 275..1637  
 US-08-348-518C-3

Alignment Scores:  
 Pred. No.: 0.208 Length: 5115  
 Score: 101.00 Matches: 34  
 Percent Similarity: 41.51% Conservative: 10  
 Best Local Similarity: 32.08% Mismatches: 30  
 Query Match: 9.70% Indels: 32  
 DB: 3 Gaps: 5

US-09-782-953-9 (1-197) x US-08-348-518C-3 (1-5115)

QY 102 ProProGlnProAlaLysGlnPheLeuLeuSerProProSerSerProProValGlyTip 121  
 DB 296 CGCGCTCAACCGGCCCCCGAGGCGCCAGGCGCGCTTCGAGCCCGCGAGGGGCGAG 355  
 QY 122 LysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaLys 141  
 DB 356 GGCCCGCGCGTCC----- 367  
 QY 142 LeuGlyProGlyGlyLysTyrGluLeuHisAlaGlyThrCluSerThrProSerVal--- 160  
 DB 368 ---GGACCCCGGGCAA-----CCGGCACCCCGCGGCGAGCCAGCGCGCGCGAGGCAACC 418  
 QY 161 -----ValValHisVal---CysAspSerAspMetGluGluGluGlu--- 173  
 DB 419 CCGCGCGGCGCATCAGATCGTCAGCTCCCGCGGGGACTCGGAGACCGGACTGGAGGGCGCTC 478  
 QY 174 -----AspProLysThrSerProLysProLysIleIleGlnThrArgArg 188  
 DB 479 TTCAACGCGCGTCATGAACCCCAAGCGGCAACGCGGCGCGCGCGCGCGCGCGCGCGCTC 538  
 QY 189 ProGlyLeuProProSer 194  
 DB 539 CGGAAGCTGCCCGACTCC 556

# RESULT 3

US-08-476-509B-3  
 Sequence 3, Application US/08476509B  
 Patent No. 6034212

GENERAL INFORMATION:  
 APPLICANT: SUDOL, MARIUS  
 APPLICANT: PEER, BORK  
 APPLICANT: HENRY, CHEN

TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
 TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
 TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTICS  
 TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,509B

FILING DATE: 01-DEC-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-101 CIP

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201 487-5800  
 TELEFAX: 201 343-1684  
 TELEX: 133521

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5115 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Human

IMMEDIATE SOURCE:

CLONE: YAP

FEATURE:

NAME/KEY: CDS

LOCATION: 275..1637

US-08-476-509B-3

Alignment Scores:

Pred. No.: 0.208 Length: 5115  
 Score: 101.00 Matches: 34  
 Percent Similarity: 41.51% Conservative: 10  
 Best Local Similarity: 32.08% Mismatches: 30  
 Query Match: 9.70% Indels: 32  
 DB: 3 Gaps: 5

US-09-782-953-9 (1-197) x US-08-476-509B-3 (1-5115)

QY 102 ProProGlnProAlaLysGlnPheLeuLeuSerProProSerSerProProValGlyTip 121  
 DB 296 CGCGCTCAACCGGCCCCCGAGGCGCCAGGCGCGCTTCGAGCCCGCGAGGGGCGAG 355  
 QY 122 LysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaLys 141  
 DB 356 GGCCCGCGCGTCC----- 367  
 QY 142 LeuGlyProGlyGlyLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal--- 160  
 DB 368 ---GGACCCCGGGCAA-----CCGGCACCCCGCGGCGAGCCAGCGCGCGCGAGGCAACC 418  
 QY 161 -----ValValHisVal---CysAspSerAspMetGluGluGluGlu--- 173  
 DB 419 CCGCGCGGCGCATCAGATCGTCAGCTCCCGCGGGGACTCGGAGACCGGACTGGAGGGCGCTC 478  
 QY 174 -----AspProLysThrSerProLysProLysIleIleGlnThrArgArg 188  
 DB 479 TTCAACGCGCGTCATGAACCCCAAGCGGCAACGCGGCGCGCGCGCGCGCGCGCGCTC 538  
 QY 189 ProGlyLeuProProSer 194  
 DB 539 CGGAAGCTGCCCGACTCC 556

# RESULT 4

US-09-103-840A-1/c

Sequence 1, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 1

```

; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 5.08e+04 Length: 4411529
Score: 89.50 Matches: 53
Percent Similarity: 35.56% Conservative: 27
Best Local Similarity: 23.56% Mismatches: 85
Query Match: 8.60% Indels: 60
DB: 4 Gaps: 10

US-09-782-953-9 (1-197) x US-09-103-840A-1 (1-4411529)
Oy 8 Cysaapvalserthrleuvalala-----Cys---Val 17
Db 4094824 TCGCGCGCTGCGAGCGCTGTGGCCCGTCGATGCCGCTCGACACCGCCAGTGCAGGCTG 4094765
Oy 18 ValaapvalgluvalpheThranglnglvallyvglylvphegluglyleuphearg 37
Db 4094764 GTGACCTCGACGCTGCTGCACCGTGCAGGTGGCTGCGCTCGCGGTGTGGCGACC 4094705
Oy 38 ThrTyrapglucyvalthrphelinleu----- 47
Db 4094704 GCCACCGCGCGGCGGCGGCCAGGTCGCCACGACACCGGTTGACGAGTTAGT 4094645
Oy 48 -----Phelyserphearpharvalargileanphe-----serHispro 61
Db 4094644 TCTATCGGCAAGCGAGATATCGGACGAGAGCTAGTGTGGGTACGGGGAGAACCA 4094585
Oy 62 Lyvseralalaargalargilegileuhtvglythglnpheargglylyslyleu 81
Db 4094584 ACTACGGCGCAGACGACCTCGGGGTGTAGTCCCGAGAACAGG---CGCGGCCCGCCACG 4094528
Oy 82 Lylleu-----Tyrpheaiglvalglnthr-Progluthraspel 95
Db 4094527 CGGCTACGTCACCCCGTCATGCCCTCTGCGCGACGACGACGCGCGG----- 4094479
Oy 95 YAspYsluhtslaleuAlaproglinproAlalyvglnpheuhtleserProprose 115
Db 4094478 -----GTGCGCGCGCGCGCCGACGTTGCCCGCGCGCGCGCGCGCCCACTCCGCG 4094435
Oy 115 rSerProprovalglutyrlypserprolleserApsalathrProvalleuanTyrAspHe 135
Db 4094434 GTTACACCGTGTGCGCGCGCGCGGACGACGCGCGCGACACCG----- 4094392
Oy 135 uLeuTyralaValaValyleuglyProgllyluyeryrglyleuhtslalegltYhrcI 155
Db 4094391 -----CCGCGCGCGCGCGCGCGCGCGCGCGCTTGTGCGCTGCGCGCGCGGATAG 4094339
Oy 155 userthrProserValaValaHisValaCysAspser-----AspMetgl 170
Db 4094338 CAGCGCGCGCGGACACCGGCTCCCGCTGTGCGCGGCTGACTACCGCGCGACCGCGC 4094279
Oy 170 uGluglugluAapProlyslleleglInthrArgProgl 190
Db 4094278 GCGCGCGCGCGCGCGCGCGCGCGCGGACGCCAACCGAGGCTT-----AACCGCGC 4094231
Oy 190 YleuProProser 194
Db 4094230 GTTCCCGCGCGC 4094218

RESULT 5
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5362224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Inail Joong, Ratna K. Vadlamudi
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR

```

```

; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REFERENCE/DOCKET NUMBER: 36,207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3157
; US-08-574-959A-8

Alignment Scores:
Pred. No.: 2.49 Length: 3211
Score: 89.00 Matches: 29
Percent Similarity: 42.86% Conservative: 13
Best Local Similarity: 29.59% Mismatches: 44
Query Match: 8.55% Indels: 12
DB: 5 Gaps: 5

US-09-782-953-9 (1-197) x US-08-574-959A-8 (1-3211)
Oy 101 AlaProProglinproAlalyeglnpheuhtleserProproserProprovalgly 120
Db 2689 GCACACACACCCGCAACCTGCTGCACGCTGCTCCCTCCCTGAGTCTCCCA---AAG 2745
Oy 121 TrplysProolleserApsalathrProvalleuanTyrAspLeuTyralaVala 140
Db 2746 GTGCAGCGAACCCGACCCGCAACCC-----GGGCTGCTTTTGGAGAGTGAG 2793
Oy 141 LylleuGlyProgllyglulyletyrglyleuhtslalegltYhrcIuserthrProserVal 160
Db 2794 GAG-----CCAGGAGCGAGAGAG-----CGTGGGCTGACACAGCTCCACCTG 2841
Oy 161 ValaValaHisValaCysAspserAspMetglu-----GluglugluAapProlythr 177
Db 2842 GCGCTGAAGCGCTCCCTCCAGGAGAGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 2901
Oy 178 SerProlysProlyslleleglInthrArgProgllyleuProProserVal 195
Db 2902 GGGCGCGCTCCCGCAGAGCTGTGAGAGAGCCCTCTCTCTCCCAACCTG 2955

RESULT 6
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Inail Joong, Ratna K. Vadlamudi
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR

```



```

Sequence 6, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Insil Jyoung, Ratna K. Vadiamudi
and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3901 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 439..3847
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-357-014-6
Alignment Scores:
Pred. No.: 3.26 Length: 3901
Score: 89.00 Matches: 29
Percent Similarity: 42.86% Conservative: 13
Best Local Similarity: 29.59% Mismatches: 44
Query Match: 8.55% Indels: 12
Gaps: 5
US-09-782-953-9 (1-197) x US-09-357-014-6 (1-3901)
QY 101 AAlaProGlnProAlaLysGlnPheLeuIleSerProIleSerSerProProValGly 120
Db |||||
QY 3379 GCACCAACACCCCAACCTGCTCCAGCTGCTGCTCCCTCCCTGATCTCCCA--AAG 3435
Db |||||
QY 121 TrpLysProIleSerAspAlaThrProValLeuAenTyAspLeuIleuTyAlaValAla 140
Db |||||
QY 3436 GTGCAGCCAGAACCCGACCCGACCC--GGGCTGCTTTGGAGTGGAG 3483
Db |||||
QY 141 LysLeuGlyProGlyGlyLysTyGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db |||||
QY 3484 GAG-----CCAGGAGCGAGAGAGAG-----CGTGGGCTGACACAGCTCCACCCCTG 3531
Db |||||
QY 161 ValValHisValCysAspSerAspMetGlu-----GluGluGluAspProLysThr 177
Db |||||
QY 3532 GCCCCTGAAGCCCTCCCTCCCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3591
Db |||||
QY 178 SerProLysProLysIleIleGlnThrArgArgProGlyLeuProProSerVal 195

```

```

Db 3592 GGGCCCCCTCCCAAGAGCTTTGTAAGAGAGCCCTCTCTCTCTCTCTCTCTCTCT 3645
RESULT 9
US-09-372-668-3
Sequence 3, Application US/09372668
Patent No. 6414129
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Jeffery, Eric W.
APPLICANT: Hjerild, Kathryn A.
APPLICANT: Ramdell, Fred
TITLE OF INVENTION: IDENTIFICATION OF THE GENE CAUSING THE
MURDER OF INVENTION: MOUSE SCURFY PHENOTYPE AND ITS HUMAN ORTHOLOG
FILE REFERENCE: 240083.501
CURRENT APPLICATION NUMBER: US/09/372,668
CURRENT FILING DATE: 1999-08-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatsSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1869
TYPE: DNA
ORGANISM: Homo sapien
US-09-372-668-3
Alignment Scores:
Pred. No.: 4.91 Length: 1869
Score: 83.50 Matches: 47
Percent Similarity: 30.56% Conservative: 19
Best Local Similarity: 21.76% Mismatches: 69
Query Match: 8.02% Indels: 81
Gaps: 7
US-09-782-953-9 (1-197) x US-09-372-668-3 (1-1869)
QY 52 ArgArgValArgIleAsnPheSerHisProLysSerAlaAla-----Arg 66
Db |||||
QY 33 AGAAGAGAGAGGTGTGCGCTTCCACACCGTACAGCGTGTCTTTCTTCTCGATATAAA 92
Db |||||
QY 67 AlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLysLeuTyPheAla 86
Db |||||
QY 93 GCAGAGTCTTTTGATACGTGACAGTTTCCACAGCCAGCAGT--ATCTTTCTGT 149
Db |||||
QY 87 GlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAla 106
Db |||||
QY 150 CAGTCCACTTACCAAGCTGCTGCTGACAAAGACCGAGTCCCAACCCAGAGCTGCG 209
Db |||||
QY 107 LysGlnPheLeuIleSerProIleSerSerProProValGlyTrpLysProIleSerAsp 126
Db |||||
QY 210 AAG-----CCCTCGGCGCTTCTCTGCGCTTGGCCCTGACATCCCAAGA 251
Db |||||
QY 127 AlaThrProValLeuAen-----TyAspLeuLeuTyAlaValAlaLys 141
Db |||||
QY 252 GCTTGGCCCACTGAGAGGCTGACACCCAAAGCTTCAGACTGTGGGGCC----- 302
Db |||||
QY 142 LeuGlyProGlyGlyLysTyGlu-----LeuHisAlaGlyThr 154
Db |||||
QY 303 CGGGGCCCAAGGGAGAACCTTCAGAGCGGAGATCTTCGAGGGGGGCCCATGCTCTCT 362
Db |||||
QY 155 GluSer----- 156
Db |||||
QY 363 TCTTCTTGAACCCCATGACATCGAGCTGAGCTGCCACATGCCCCCTACTCATG 422
Db |||||
QY 156 ----- 156
Db |||||
QY 423 GTGCACCTCCGGGAGCAGGCTGGGCCCTTGGCCCATTAAGAGCATCTCTCCAGAG 482
Db |||||
QY 157 -----ThrProSerVal 160
Db |||||
QY 483 AGGCACATTTTCATGACACAGCTCTCAAGGTGATGCCACGCCGAGACCCCTGTGTG 542
Db |||||
QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
Db |||||

```

```

Db 543 CAGGTGACCCCTCGAGAGCCCGCCATGATCAGCCTCACACCCACCCAGCCACT 602
Qy 181 ProlysillelleGlnThrArgArgProGlyLeuProSerValSer 196
Db 603 GGGGTCTTCTCCCTCAAGGCCCGGCTGGCTCCACCTGGGATCAAC 650

RESULT 10
US-08-808-599A-23/c
; Sequence 23, Application US/08080599A
; Patent No. 611089
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin, Trophinin-Assisting
; TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,599A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3483 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3481
; US-08-808-599A-23

Alignment Scores:
Pred. No.: 15 2 Length: 3483
Score: 82.50 Matches: 31
Percent Similarity: 41.75% Conservative: 12
Best Local Similarity: 30.10% Mismatches: 49
Query Match: 7.93% Indels: 11
DB: 3 Gaps: 4

US-09-782-953-9 (1-197) x US-08-808-599A-23 (1-3483)
Qy 100 LeuAlaProGlnProLalysGlnPheLeuIleSerProProSerSerProProVal 119
Db 3164 CTTGGTCCACCGAAGCCAGGTCTCTCTCTGGTCCACCGAAGCCAGGTG 3111
Qy 120 GlyTrpIleProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaVal 139
Db 3110 ---CTTGGTCCACTCGAGAGCCAGCAGCAGGTCTGGTTCACCACTAAAGCCAACTA 3054

```

```

QY 140 AlaLysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSer 159
      :::::|||||
Db 3053 ---GAACCTGGGTCATTACTGAAGCTGGAGCCAGTGTGGGTTCCACCATTAAGCCATTG 2997
      :::::|||||
QY 160 ValValValHisValCysAspSerAspMet-----GluGluGlu 172
      :::::|||||
Db 2996 CTGGTGTTGGAAACCACTGCCAAGCAATGATGGTATTAGTCTCTGTTGAAGCTGGCA 2937
      :::::|||||
QY 173 GluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192
      |||||
Db 2936 TTAGGCCCATTCGCAAGCAATGCCATTGTTGAGCTGTACACTAAAGCTGCACCATGTG 2877
      |||||

QY 193 ProSerVal 195
      |||||
Db 2876 CCAAGTGTG 2868

RESULT 11
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 2,71e+05 Length: 4411529
Score: 82.50 Matches: 31
Percent Similarity: 43.48% Conservative: 9
Best Local Similarity: 33.70% Mismatches: 32
Query Match: 7.93% Indels: 20
DB: 4 Gaps: 5

US-09-782-953-9 (1-197) x US-09-103-840A-1 (1-4411529)
QY 50 SerPheArgArgVal-----ArgIleAsnPheSerHisProLysSerAlaAlaArgala 67
      |||||
Db 3667235 TCGTCTTGAGGTCAACAGCGCGCTTCAGGTAGAACACCCGGTCCCGAGGAAACCGCG 3667294
      |||||
QY 68 ArgIleGluLeuHisGluThrGlnPheArg-----GlyLysLysLeuLysLeuTyr 84
      |||||
Db 3667295 GGCATGACCTTGGTGTGTGTCAGCAATTCGGATCGCCAAACGCGGAAAGCTGGACATC--- 3667351
      |||||
QY 85 PheAlaGlnValGlnThrProGlu-----ThrAsp 94
      :::::
Db 3667352 ---ACCGAGGATCCCAACCCCGCGCGGACGCCATCGAATTCGGATCAACGCGGAGGAC 3667408
      |||||
QY 95 GlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeuIleSerProPro 114
      :::::
Db 3667409 GCGGGGGGTAACTTCTTACGGCGCCCGGGCCGGTGACAAGTTC-----CACCCGCGC 3667462
      |||||

QY 115 SerSerProProVal-GlyTrpLysProIleSer 125
      |||||
Db 3667463 TCCGGGCCCGGTGTGGGTGGACTCCGGTGTGCG 3667496
      |||||

RESULT 12
US-08-340-203A-2
; Sequence 2, Application US/08340203A
; Patent No. 5756668

```

Db	3263	CAC	TAGCCCAAGGCGCAACCTTCTCTCTCAGTCA	CCAAAGGCGGAGATTCGAGTCG	3322	183
Qy	168	AspMetGluGlnJugJubAppProlyerThr	---	SerProlySProlySile	-----	183
Db	3323	-----	GAAAGCGAAGAGCTTCCACAGAGTCTCCCACTCCCGGAGTCCCTCCCTT	3373		
Qy	184	-----	11egInThrArgArgProGlyLeuProProSer	194		
Db	3374	CCCTTCCTCGGCGCCCGAGACCAATTTATTCAGATCGCCCGCGGCGCCCAATTC	3430			
	RESULT 13					
	US-08-452-567-2					
	Sequence 2, Application US/08452567					
	Patent No. 5846712					
	GENERAL INFORMATION:					
	APPLICANT: Baylin, Stephen B.					
	APPLICANT: Wales, Michele M.					
	TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1					
	NUMBER OF SEQUENCES: 3					
	CORRESPONDENCE ADDRESS:					
	ADDRESSEE: Fish & Richardson P.C.					
	STREET: 4225 Executive Square, Suite 1400					
	CITY: La Jolla					
	STATE: California					
	COUNTRY: USA					
	ZIP: 92037					
	COMPUTER READABLE FORM:					
	MEDIUM TYPE: Floppy disk					
	COMPUTER: IBM PC compatible					
	OPERATING SYSTEM: PC-DOS/MS-DOS					
	SOFTWARE: PatentIn Release #1.0, Version #1.25					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/452,567					
	FILING DATE: 25-MAY-1995					
	CLASSIFICATION: 435					
	ATTORNEY/AGENT INFORMATION:					
	NAME: Halle, Ph.D., Lisa A.					
	REGISTRATION NUMBER: 38,347					
	REFERENCE/DOCKET NUMBER: 07265/039003 (Div PD3664)					
	TELECOMMUNICATION INFORMATION:					
	TELEPHONE: (619) 678-5070					
	TELEFAX: (619) 678-5099					
	INFORMATION FOR SEQ ID NO: 2:					
	SEQUENCE CHARACTERISTICS:					
	LENGTH: 4112 base pairs					
	TYPE: nucleic acid					
	STRANDEDNESS: single					
	TOPOLOGY: linear					
	MOLECULE TYPE: DNA (genomic)					
	IMMEDIATE SOURCE:					
	CLONE: HIC-1 coding polynucleotide					
	FEATURE:					
	NAME/KEY: CDS					
	LOCATION: 1086..2726					
	US-08-452-567-2					
	Alignment Scores:					
	Pred. No.:	21.8		Length:	4112	
	Score:	82.00		Matches:	41	
	Percent Similarity:	33.96%		Conservative:	13	
	Best Local Similarity:	25.79%		Mismatches:	41	
	Query Match:	7.88%		Indels:	64	
	DB:	2		Gaps:	9	
	US-09-782-953-9 (1-197) x US-08-452-567-2 (1-4112)					
Qy	93	ThraSpGlyAspRlyLeuHisLeuAlaProGlnProGlnProAlaGlnInPheLeuIleSer	112			
Db	2975	ACCAAAAGAGAGCCCCAGGCCCTC--CGGCTCTTCTGTGTGTTGCT--CGG	3022			
Qy	113	ProProSerSerProPro--	118			



```
Db 3028 CCCCCCTCCCCGGCTCCGGCTGCTCTTAGAGGGGAGGGGTGTCACTGTTCGGGGCACTC 3082
Qy 119 ValGlyTrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAla 138
Db 3083 CTAGCCCTACCTCCGGCCCTCCGACACACCCATTCTCACTGTGAATCTCCCGCTGGG 3142
Qy 139 Val-----AlaLysLeuGlyProGlyGluLysTyrGlu----- 149
Db 3143 TCGGAGCGTTCGGGAGAGTGGGGAGTGGGGAGGAGTGAAGCCGGCGGAGGCCCGCC 3202
Qy 150 -----LeuHisAlaGlyThrGluSerThrProSerValValValHis----- 163
Db 3203 ACCCCCGCTCCACCCACCCCGGAGTCAATGTGAAGTTCCTCATTTTGCAAGTGG 3262
Qy 164 -----ValCysAspSer 167
Db 3263 CACTAGCCAGGCGCAACCTTCCTCAGTCACCAAGGCGGGAGTTCGGAGTGG 3322
Qy 168 AspMetGluGluGluAspProLysThr---SerProLysProLysIle----- 183
Db 3323 -----GAAGCGAAGAGCTACACAGGCTCTCCACTCCCGGGGTGCTCCCTT 3373
Qy 184 -----IleGlnThrArgArgProGlyLeuProProSer 194
Db 3374 CCCTTCCTCGGGCCCGGACCATATTTATTGTCATGCGCCCGGGCCCGCATCC 3430

RESULT 14
US-08-452-427-2
; Sequence 2, Application US/08452427
; Patent No. 5922590
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/340,203
; FILING DATE: 15-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: 07265/039001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HIC-1 coding polynucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1086...2726
```

```
US-08-452-427-2
Alignment Scores:
Pred. No.: 21.8 Length: 4112
Score: 82.00 Matches: 41
Percent Similarity: 33.96% Conservative: 13
Best Local Similarity: 25.79% Mismatches: 41
Query Match: 7.88% Indels: 64
DB: 2 Gaps: 9

US-09-782-953-9 (1-197) x US-08-452-427-2 (1-4112)
Qy 93 ThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeuIleSer 112
Db 2975 ACCAAGGAGACCCAGGCCCTC-----CGCCCTCTTCCTGTGTTCTGTT 3022
Qy 113 ProProSerSerProPro----- 118
Db 3023 CCCCCCTCCCGGCTCCGGCTGCTCTTAGAGGGGAGGGGTGTCACTGTTCGGGGCACTC 3082
Qy 119 ValGlyTrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAla 138
Db 3083 CTAGCCCTACCTCCGGCCCTTCGGACACACCCATTCTCACTGTGAATCTCCCGCTGGG 3142
Qy 139 Val-----AlaLysLeuGlyProGlyGluLysTyrGlu----- 149
Db 3143 TCGGAGCGTTCGGGAGAGTGGGGAGTGGGGAGGAGTGAAGCCGGCGGAGGCCCGCC 3202
Qy 150 -----LeuHisAlaGlyThrGluSerThrProSerValValValHis----- 163
Db 3203 ACCCCCGCTCCACCCACCCCGGAGTCAATGTGAAGTTCCTCATTTTGCAAGTGG 3262
Qy 164 -----ValCysAspSer 167
Db 3263 CACTAGCCAGGCGCAACCTTCCTCAGTCACCAAGGCGGGAGTTCGGAGTGG 3322
Qy 168 AspMetGluGluGluAspProLysThr---SerProLysProLysIle----- 183
Db 3323 -----GAAGCGAAGAGCTACACAGGCTCTCCACTCCCGGGGTGCTCCCTT 3373
Qy 184 -----IleGlnThrArgArgProGlyLeuProProSer 194
Db 3374 CCCTTCCTCGGGCCCGGACCATATTTATTGTCATGCGCCCGGGCCCGCATCC 3430

RESULT 15
US-09-085-407-2
; Sequence 2, Application US/09085407
; Patent No. 6103877
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,407
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,203
; FILING DATE: 15-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME/KEY: CDS
; LOCATION: 1086...2726
```

NAME: Haile, Ph.D., Lisa A.  
REGISTRATION NUMBER: P-38,347  
REFERENCE/DOCKET NUMBER: 07265/039001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: HIC-1 coding polynucleotide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1086..2726  
US-09-085-407-2

Alignment Scores:  
Pred. No.: 21.8 Length: 4112  
Score: 82.00 Matches: 41  
Percent Similarity: 33.96% Conservative: 13  
Best Local Similarity: 25.79% Mismatches: 41  
Query Match: 7.88% Indels: 64  
Gaps: 9  
DB:

US-09-782-953-9 (1-197) X US-09-085-407-2 (1-4112)

QY 93 ThrAspGlyAspIysLeuHisLeuAlaProProGlnProAlaIysGlnPheLeuIleSer 112  
DB 2975 ACCAAGAGACACCCAGGCCCTC---CGGCTCTTCCTGTGTTGCT-----CGG 3022  
QY 113 ProProSerSerProPro----- 118  
DB 3023 CCCCCTCCCGGCTCGGCTGCTTAGAGGGAGGGGTGTCACGTGCGGGCACTC 3082  
QY 119 ValGlyTyrIlyProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAla 138  
DB 3083 CTAGCCCTACCTCCGGCCCTTGCGACCACTTCTCAGTGAATCTCCCGGCTGGG 3142  
QY 139 Val-----AlaIysLeuGlyProGlyGlnIlyTyrGlu----- 149  
DB 3143 TCGGAGCGTGGGAGAGTGGGAGTGGGAGGAGTGAAGCCGCGCGAGGCCCGCC 3202  
QY 150 -----LeuHisAlaGlyThrGluSerThrProSerValValValHis----- 163  
DB 3203 ACCCCGCTCCACCCACCCCGGAGTGAATGTGAAGTTCTCATTTGCAAGAGTGG 3262  
QY 164 -----ValCysAspSer 167  
DB 3263 CACTAGCCCAAGGCCAACCTTCTCTCAGTACCAAGGCGGAGGAGTTCTGAGTGG 3322  
QY 168 AspMetGluGluGluAspProIlyThr---SerProIlyProIlySile----- 183  
DB 3323 -----GAAAGCGAAGAGCTTACCACTCCCGGCTGCGCTCCCTT 3373  
QY 184 -----IleGlnThrArgProGlyLeuProProSer 194  
DB 3374 CCTTCCCTGGGCGCCGGAACATATTATTGCAATGCGCCCGGCGGCCCAATCC 3430

Search completed: December 14, 2002, 22:55:07  
Job time : 1511.9 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 17:10:31 ; Search time 20:3119 Seconds  
(without alignment)  
11623.258 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggattttgggacttag.....caccgatccacttagctga 597

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	100.0	597	10	US-09-782-953-5
2	515.8	86.4	597	10	US-09-782-953-2
3	504.4	84.5	599	10	US-09-782-953-1
4	470.6	78.8	2331	10	US-09-782-953-11
5	470.6	78.8	2355	12	US-10-044-090-255
6	408.2	68.4	2212	10	US-09-782-953-14
7	399.6	66.9	2173	10	US-09-880-107-3340
8	395.4	66.2	2358	10	US-09-925-302-347
9	214.2	35.9	3184	10	US-09-954-456-497
10	214.2	35.9	3184	10	US-09-782-953-17
11	201.4	33.7	594	10	US-09-782-953-8
12	191.2	32.0	720	10	US-09-782-953-23
13	191.2	32.0	828	10	US-09-782-953-20
14	144.6	24.2	412	10	US-09-864-761-10388
15	144.6	24.2	446	10	US-09-864-761-2064
16	123	20.6	486	10	US-09-864-761-809
17	91.8	15.4	111	10	US-09-864-761-27019
18	84.4	14.1	123	10	US-09-864-761-17592
19	81.4	13.6	365	10	US-09-728-445-736

C	20	70.6	11.8	85	10	US-09-864-761-18808	Sequence 18808, A
	21	37.8	6.3	2853	10	US-09-965-631-3	Sequence 3, Appli
	22	37.8	6.3	3446	10	US-09-965-631-7	Sequence 7, Appli
	23	36.2	6.1	5739	10	US-09-960-253-142	Sequence 142, App
	24	36.2	6.1	6417	10	US-09-962-436-288	Sequence 288, App
	25	35.6	6.0	490	10	US-09-880-107-2261	Sequence 2261, Ap
	26	34.4	5.8	497	10	US-09-822-263-35	Sequence 35, Appl
	27	34.4	5.8	4363	10	US-09-764-864-241	Sequence 17, Appl
	28	34	5.7	345	10	US-09-822-263-17	Sequence 241, App
	29	34	5.7	387	10	US-09-920-300A-395	Sequence 395, App
	30	34	5.7	387	12	US-10-033-528-395	Sequence 395, App
	31	34	5.7	438	9	US-09-736-457-1028	Sequence 1028, Ap
	32	34	5.7	438	9	US-09-902-941-1028	Sequence 1028, Ap
	33	34	5.7	497	10	US-09-822-263-33	Sequence 33, Appl
	34	34	5.7	1147	10	US-09-880-107-2311	Sequence 2311, Ap
	35	33.8	5.7	156	10	US-09-864-761-20014	Sequence 20014, A
	36	33.8	5.7	415	10	US-09-864-761-3236	Sequence 3236, Ap
	37	33.6	5.6	1752	9	US-09-887-552A-1	Sequence 1, Appli
	38	33.4	5.6	565	10	US-09-962-436-116	Sequence 116, App
	39	33.4	5.6	565	10	US-09-954-456-1735	Sequence 1735, Ap
	40	33.4	5.6	575	10	US-09-864-761-8864	Sequence 8864, Ap
	41	33.4	5.6	20247	10	US-09-764-877-2680	Sequence 2680, Ap
	42	33.2	5.6	299	10	US-09-864-761-21553	Sequence 21553, A
	43	33.2	5.6	470	10	US-09-864-761-2442	Sequence 2442, Ap
	44	33.2	5.6	1875	9	US-10-001-835-55	Sequence 56, Appl
	45	33.2	5.6	2289	10	US-09-850-964-3	Sequence 3, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-782-953-5

; Sequence 5, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
CALCINEURIN INTERACTING PROTEIN (NCIP)

; FILE REFERENCE: UTSD:674PZ1

; CURRENT APPLICATION NUMBER: US/09/782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentin ver. 2.1

; SEQ ID NO 5

; LENGTH: 597

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(594)

; US-09-782-953-5

##### Query Match

Best Local Similarity 100.0%; Score 597; DB 10; Length 597;

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTTAGGGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGGCAACGAT 60

Db 1 ATGGATTTAGGGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGGCAACGAT 60

QY 61 GATGCTTTCACGGAAGTGACCCAGGGCCAAATTTGAATCCCTTTTCAAGCAATATGAC 120

Db 61 GATGCTTTCACGGAAGTGACCCAGGGCCAAATTTGAATCCCTTTTCAAGCAATATGAC 120

QY 121 AAGGACACCACTTCCAGTAGTTTTTAAGAGCTTCAAGCTGTCGGATTAACACTTCAGCAAC 180

Db 121 AAGGACACCACTTCCAGTAGTTTTTAAGAGCTTCAAGCTGTCGGATTAACACTTCAGCAAC 180

QY 181 CCCTTATCTGACGCCGATGCCAGCTCGGCTGCAACAGCCGAGTTCTCTGGGAGGAA 240

```

|||
Db 181 CCTTATGAGCCGATGCGAGGTGCGGCTGCACAAAGCCGAGTTCCTGGGAAAGAA 240
Qy 241 ATGAAGTTGATTTTGGCTCAGACTTTACATATAGAAAGTTGACACCTGGCTCCGCCAAT 300
Db 241 ATGAAGTTGATTTTGGCTCAGACTTTACATATAGAAAGTTGACACCTGGCTCCGCCAAT 300
Qy 301 CCCGACAAACAGTTCTCATCTCTCCGCTCTCTCCGCTGGTGAACAAAGTA 360
Db 301 CCCGACAAACAGTTCTCATCTCTCCGCTCTCTCCGCTGGTGAACAAAGTA 360
Qy 361 GAGATGACACCCCGCATATAATAGCATCTTTATATGTCATCTCCAGCTGGGCGCA 420
Db 361 GAGATGACACCCCGCATATAATAGCATCTTTATATGTCATCTCCAGCTGGGCGCA 420
Qy 421 GGAGAGAGATGAACTGCATGACGACGACAGCCCATCTCCAGTGTGTGTCACAGTG 480
Db 421 GGAGAGAGATGAACTGCATGACGACGACAGCCCATCTCCAGTGTGTGTCACAGTG 480
Qy 481 TGTGAGAGTGAACCAAGATGAGAGAGAGAGAGATGAGAGATGAGAGAGAGAGAGAG 540
Db 481 TGTGAGAGTGAACCAAGATGAGAGAGAGAGAGATGAGAGATGAGAGAGAGAGAGAG 540
Qy 541 AAGCCCAATATCATCGACGACGAGAGCCGAGTACACACCGATCCACCTTAGCTGA 597
Db 541 AAGCCCAATATCATCGACGACGAGAGCCGAGTACACACCGATCCACCTTAGCTGA 597

```

RESULT 2  
US-09-782-953-2  
Sequence 2, Application US/09782953  
Patent No. US20020150953A1

```

GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: US/09/782,953
CURRENT APPLICATION NUMBER: US/09/782,953
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-09-782-953-2

```

Query Match 86.4%; Score 515.8; DB 10; Length 597;  
Best Local Similarity 94.4%; Pred. No. 2,3e-151;  
Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

Qy 31 AGCTCCCTGATTTGCTGTGTGTCGCAACGATGATGCTTTCAGCGAAGTGAAGACGAGGCC 90
Db 31 AGCTCCCTGATTTGCTGTGTGTCGCAACGATGATGCTTTCAGCGAAGTGAAGACGAGGCC 90
Qy 91 AAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTAAGAGC 150
Db 91 AAATTTGAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTAAGAGC 150
Qy 151 TTCAAAAGTGTCCGATTAATCTTGAAGCAACCCCTTATCTGAGCGGATGCGAGGCTCGG 210
Db 151 TTCAAAAGTGTCCGATTAATCTTGAAGCAACCCCTTATCTGAGCGGATGCGAGGCTCGG 210
Qy 211 CTGACAAAGACGAGTCTGCTGGGGAAGAAATGAAGTTGATTTTGGCTCAGACTTTACAC 270
Db 211 CTGACAAAGACGAGTCTGCTGGGGAAGAAATGAAGTTGATTTTGGCTCAGACTTTACAC 270

```

```

Qy 271 ATGAGAGTTCAACCTGCTCCGCCCAATCCCGCAACAAAGTTTCTCATCTCCCTCG 330
Db 271 ATGAGAGTTCAACCTGCTCCGCCCAATCCCGCAACAAAGTTTCTCATCTCCCTCG 330
Qy 331 GCCTTCCTCCCGTGGTGGTGAACAAAGTGAAGATGCAACCCCGTCAATAATAGAT 390
Db 331 GCCTTCCTCCCGTGGTGGTGAACAAAGTGAAGATGCAACCCCGTCAATAATAGAT 390
Qy 391 CTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGATGAACTGATGACGACA 450
Db 391 CTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGATGAACTGATGACGACA 450
Qy 451 GACCCCACTCCAGTGTGTGTGTCACGCTGTGAGAGAGTGAACCAAGAGAAATGAGAGAA 510
Db 451 GACCCCACTCCAGTGTGTGTGTCACGCTGTGAGAGAGTGAACCAAGAGAAATGAGAGAA 510
Qy 511 GAGAGAGATGAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Db 511 GAGAGAGATGAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Qy 571 GAGTACACACCGATCCACCTTAGCTGA 597
Db 571 GAGTACACACCGATCCACCTTAGCTGA 597

```

RESULT 3  
US-09-782-953-1  
Sequence 1, Application US/09782953  
Patent No. US20020150953A1

```

GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: US/09/782,953
CURRENT APPLICATION NUMBER: US/09/782,953
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 599
TYPE: DNA
ORGANISM: Mus musculus
US-09-782-953-1

```

Query Match 84.5%; Score 504.4; DB 10; Length 599;  
Best Local Similarity 99.6%; Pred. No. 8,4e-148;  
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

Qy 4 GATTTTGAAGACTTTAGCTTACAAATTTTAACTGCTCTGATTCCTTGTGTGTCGCAAGATGAT 63
Db 83 GATTTTGAAGACTTTAGCTTACAAATTTTAACTGCTCTGATTCCTTGTGTGTCGCAAGATGAT 142
Qy 64 GCTTCAGGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Db 143 GCTTCAGGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
Qy 124 GACACCACTTCAGATTTTAAAGCTTCAACGCTGTCCGATTAATCTTGAAGCAACCC 183
Db 203 GACACCACTTCAGATTTTAAAGCTTCAACGCTGTCCGATTAATCTTGAAGCAACCC 262
Qy 184 TTATCTGACGCGGATGCCAGGCTGGGCTGACAAAGACGAGTCTGCGGGAAGAAATG 243
Db 263 TTATCTGACGCGGATGCCAGGCTGGGCTGACAAAGACGAGTCTGCGGGAAGAAATG 322
Qy 244 AAGTTATTTTGTCTCAGACTTTACATATAGAAATGCAACCTGCTCCGCCAATCC 303
Db 323 AAGTTATTTTGTCTCAGACTTTACATATAGAAATGCAACCTGCTCCGCCAATCC 381
Qy 304 GACAAACAGTTCTCATCTCCCTCGGCTCTCTCCGTTGGCTGGAACAAATAGAA 363

```

Db 382 GACAAACAGTTCTCTCATCTCCCTCCGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAA 441  
Qy 364 GATGCCACCCCGTCATAAATAGATCTTTTATATGCAATCTCCAAAGCTGGGCGCCAGGA 423  
Db 442 GATGCCACCCCGTCATAAATAGCATCTTTTATATGCAATCTCCAAAGCTGGGCGCCAGGA 501  
Qy 424 GAGAAGTATGAATGTCATGACGACGACAGACCCCACTCCCACTGGTGGTCCAGCTGTGT 483  
Db 502 GAGAAGTATGAATGTCATGACGACGACAGACCCCACTCCCACTGGTGGTCCAGCTGTGT 561  
Qy 484 GAGAGTACCAAGAGAAATGAGGAGGAGAGAGAGAGAT 521  
Db 562 GAGAGTACCAAGAGAAATGAGGAGGAGAGAGAGAT 599

RESULT 4  
US-09-782-953-11  
; Sequence 11, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: CALCIENEURIN INTERACTING PROTEIN (MCIP)  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (144)..(734)  
US-09-782-953-11

Query Match 78.8%; Score 470.6; DB 10; Length 2331;  
Best Local Similarity 87.9%; Pred. No. 6.6e-137;  
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;  
Qy 1 ATGCAATTTAGGACCTTACCTACAATTTTAGCTCCCTGATTCCTGTGTGGGCAACGAT 60  
Db 144 ATGCAATTTAGGAACTTTAACTACAGTTTATAGCTCCCTGATTCCTGTGTGGGCAACGAT 203  
Qy 61 GATGCTCTCAGCGAAAGTGAGACGAGGCGCAAAATTTGAATCCCTCTTCAGACATATGAC 120  
Db 204 GATATCTTCAGCGAAGTGAAACGAGGCGCAAAATTTAGTCCCTCTTTAGGACGTATGAC 263  
Qy 121 AAGGACACACCTTCCAGTATTTAAGAGCTTCAAAAGCTGTCGGGATAAACTTCAGCAAC 180  
Db 264 AAGGACATCACCTTTCAGTATTTAAGAGCTTCAAAAGAGTCAAGATTAATTCAGCAAC 323  
Qy 181 CCCTATCTGACCGCATGCCAGCTGCGGTGCAACAGACGAGTTCCTGGGGAAGGAA 240  
Db 324 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGTCATAAGACTGAGTTTCTGGGAAGGAA 383  
Qy 241 ATGAAGTTGATTTTCTCAGACTTTTACATAGGAAGTTTCAACCTGGCTCCGCCAAT 300  
Db 384 ATGAAGTTATATTTGCTCAGACTTTACATAGGAAGTTCACCTGGCTCCGCCAAT 443  
Qy 301 CCGGACAAACAGTTCTCAATCTCCCTCCGCGCTCTCTCCGCTTGGCTGGGAAACAAGTA 360  
Db 444 CCAGACAAGCAGTTCTGATCTCTCCCTCCGCGCTCTCCGCGAGTGGGATGGAAACAAGT 503  
Qy 361 GAAGATGCCACCCCGTCATAAATAGCATCTTTATATGCCATCTTCAAGCTGGGCGCA 420  
Db 504 GAAGATGCCACCCCGTCATAACTATGATCTCTTATATGCCATCTTCAAGCTGGGCGCA 563  
Qy 421 GGAGAGAAGTATGAACTGCAATGCGGACAGACCCCACTCCCACTGGTGTGTGTCCACGTG 480

Db 564 GGGGAAAAGTATGAATTCGACGAGGAGCTGACACCACTCCCACTGGTGGTCCATGTA 623  
Qy 481 TGTGAGAGTGACCAAGAGAGATGAGGAGGAGAGAGATGAGAGAGATGAGAGACCC 540  
Db 624 TGTGAGAGTGATCAAGAGAAAGGAGGAGGAA---ATGAAAAGAAATGAGGAGACCT 680  
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTCA 597  
Db 681 AAGCCCAAAATCATCCAGACGAGGAGGCGGAGTACACCGCATCCCACTCAGCTGA 737  
RESULT 5  
US-10-044-090-255  
; Sequence 255, Application US/10044090  
; Patent No. US20020137081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 255  
; LENGTH: 2355  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1  
US-10-044-090-255

Query Match 78.8%; Score 470.6; DB 12; Length 2355;  
Best Local Similarity 87.9%; Pred. No. 6.6e-137;  
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;  
Qy 1 ATGGAATTTAGGACTTTTAGCTACAAATTTAGCTCCCTGATTTAGCTTGTGTGGCAACGAT 60  
Db 151 ATGCAATTTAGAACTTTAACTACAGTTTATAGCTCCCTGATTCCTGTGTGGCAACGAT 210  
Qy 61 GATGCTTTCAGGAAAGTGAGACGAGGCGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120  
Db 211 GATATCTTTCAGGAAAGTGAAACGAGGCGCAAAATTTGAGTCCCTCTTTAGGACGTATGAC 270  
Qy 121 AAGGACACACCTTCCAGTATTTAAGAGCTTCAAAAGTTCGCGATTAACCTTCAGCAAC 180  
Db 271 AAGGACATCACCTTTTCAGTATTTAAGAGCTTCAAAAGAGTCAAGATTAACCTTCAGCAAC 330  
Qy 181 CCCTATCTGACGCGAGTCCAGGCTGCGGCTGCACAGACCGAGTTCTCTGGGAAGGAA 240  
Db 331 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 390  
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTGGCTCCGCCAAT 300  
Db 391 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGGAAGTTCACACCTGGCTCCGCCAAT 450  
Qy 301 CCGGACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAACAAGTA 360  
Db 451 CCGACAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGGATGGAAACAAGTG 510  
Qy 361 GAAGATGCCACCCCGTCATAAATTAAGATCTTTTATATGCCATCTTCAAGCTGGGCGCA 420  
Db 511 GAAGATGCCACCCCGTCATAACTATGATCTTATATGCCATCTTCAAGCTGGGCGCA 570  
Qy 421 GGAGAGAAGTATGAATGTCATGCGAGGACAGACCCCACTCCCACTGGTGGTCCACGTG 480  
Db 571 GGGGAAAAGTATGAATTCACGCGAGGACTGACACCACTCCCACTGGTGGTCCATGTA 630  
Qy 481 TGTGAGAGTGACCAAGAGAAATGAGGAGGAGAGAGATGAGAGAAATGAAGAGACCC 540  
Db 631 TGTGAGAGTGATCAAGAGAAAGGAGGAGGAA---ATGAAAAGAAATGAGGAGACCT 687

Qy 541 AAGCCAAATATCATCCAGACGAGACCGAGTACACACCGATCACTTAGCTGA 597  
 Db 668 AAGCCAAATATATCCAGACGAGACCGAGTACACACCGATCACTTAGCTGA 744

## RESULT 6

US-09-782-953-14  
 ; Sequence 14, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSID:674P21  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 2212  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (25)..(615)  
 ; US-09-782-953-14

## Query Match

Best Local Similarity 68.4%; Score 408.2; DB 10; Length 2212;  
 Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

Qy 1 ATGATTTTGAAGGAGCTTGAAGTAAATTTAGTCCCTGATGCTGTGTGGCAACGAT 60  
 Db 25 ATGAGAGAGGTGAGCTCGAGACCTGCCAGCGCCATGCTGCTGATCACTGAGCCG 84  
 Qy 61 GATGCTTCAGCAAGAAAGTAGAGACGAGGCAATTTGAAATCCCTCTTCAAGATATGAC 120  
 Db 85 CCGGTCTGTGAGCGCTGCTGCGGCGCAATTTGAGTCCCTCTTGAAGATATGAC 144  
 Qy 121 AAGGACACCACTTCAGATATTTAAGAGCTTCAAAAGTGTCCGATTAACCTTCAAGCAAC 180  
 Db 145 AAGGACATACCTTTAGATATTTAAGAGCTTCAAAAGTGTCAAGATTAACCTTCAAGCAAC 204  
 Qy 181 CCGTATTCGACCGAGTGCAGGCTGCGGCTGCAAAAGCAAGTTCCTGGGGAAGAA 240  
 Db 205 CCGTCTCCGACAGATGCGAGCTTCAGCTGATAGACTGAGTTTCTGGGAAAGAA 264  
 Qy 241 ATGAGTGTATTTTGTGAGCTTACACATAGAAAGTTCAGACCTGCTCGCCCAAT 300  
 Db 265 ATGAGTATATTTTGTGAGCTTACACATAGAAAGTTCAGACCTGCTCGCCCAAT 324  
 Qy 301 CCGGACCAACAGTTCCTCATCTCCCTCGGCTCTTCTCCGTTGGTGGAAACAAGTA 360  
 Db 325 CCGAGACAGAGTTTGTGATCTCCCTCCGCTCTCCGACAGTGGGATGAAACAAGTG 384  
 Qy 361 GAAAGTCCACCCCGCTCATTAATTAAGATCTTTTAATGCAATTCCTCAAGCTGGGGCCA 420  
 Db 385 GAAAGTCCGACCCCGCTCATTAATTAAGATCTTTTAATGCAATTCCTCAAGCTGGGGCCA 444  
 Qy 421 GGAAGAGATGATGAAGTGCAGGACAGACCCCACTCCCAAGTGTGTGCTCAAGTG 480  
 Db 445 GGGGAAAGATGATGATGACGACGACCTGACACCACTCCCAAGTGTGTGCTCAAGTG 504  
 Qy 481 TGTGAGATGACCAAGAAATGAGAGAAAGAGAGATGAGAGATGAAAGAGAGCC 540  
 Db 505 TGTGAGATGATCAAGAAAGAGAGAAAGAGAGAA--ATGAAAGAAATGAGAGAGACT 561  
 Qy 541 AAGCCAAATATCATCCAGACGAGACCGAGTACACACCGATCACTTAGCTGA 597  
 Db 562 AAGCCAAATATATCCAGACGAGACCGAGTACACACCGATCACTTAGCTGA 618

## RESULT 7

US-09-880-107-3340  
 ; Sequence 3340, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/211,379  
 ; PRIOR FILING DATE: 2000-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/237,054  
 ; PRIOR FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 3950  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3340  
 ; LENGTH: 2173  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833  
 ; US-09-880-107-3340

## Query Match

Best Local Similarity 66.9%; Score 399.6; DB 10; Length 2173;  
 Matches 448; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

Qy 88 GCCAAATTTGATCCCTCTTCAAGAAATATGACAGACACCACTTCCATATTTTAAG 147  
 Db 58 GCCAAATTTGATCCCTCTTCAAGAAATATGACAGACACCACTTCCATATTTTAAG 117  
 Qy 148 AGCTTCAAGCTGTCCGATTAACCTTCAAGCAACCTTTATTCAGACCGATGCCAGCTG 207  
 Db 118 AGCTTCAAGCTGTCCGATTAACCTTCAAGCAACCTTTATTCAGACCGATGCCAGCTG 177  
 Qy 208 CCGGTGCAAGACCGAATCTCTGCGGAGGAAATGAAGTTATTTTGTGACACTTA 267  
 Db 178 CAGGTGCAAGACCGAATCTCTGCGGAGGAAATGAAGTTATTTTGTGACACTTA 237  
 Qy 268 CACATAGGAAGTTCACACCTGCTCCGCCAATCCGCAAAAGTTCCTCATCTCCCT 327  
 Db 238 CACATAGGAAGTTCACACCTGCTCCGCCAATCCGCAAAAGTTCCTCATCTCCCT 297  
 Qy 328 CCGGCTCTCTCCCTGCTGCAACCAATGAAATGCCACCCCTCATTAATTAAC 387  
 Db 298 CCGGCTCTCTCCCTGCTGCAACCAATGAAATGCCACCCCTCATTAATTAAC 357  
 Qy 388 GATCTTTATATGCAATCCCAAGCTGGGGCCAGGAGAAATGATGATGATGATGATGAT 447  
 Db 358 GATCTTTATATGCAATCCCAAGCTGGGGCCAGGAGAAATGATGATGATGATGATGATGAT 417  
 Qy 448 ACAGACCCCACTCCAGTGTGTGCTCCAGCTGTGTGAGATGACCAAGAAATGAGAG 507  
 Db 418 ACTGACACCACTCCAGTGTGTGCTCCAGCTGTGTGAGATGACCAAGAAATGAGAG 477  
 Qy 508 GAAGAGAAAGATGAGAGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAGAAAT 567  
 Db 478 GAAGAGAA--ATGAGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAGAAAT 534  
 Qy 568 CCGAGTACACACCGATCACTTAGCTGA 597  
 Db 535 CCGAGTACACACCGATCACTTAGCTGA 564

## RESULT 8

US-09-925-302-347  
 ; Sequence 347, Application US/09925302

```
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 347
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-302-347

Query Match      66.2%; Score 395.4; DB 10; Length 2358;
Best Local Similarity 80.9%; Pred. No. 2,1e-113;
Matches 483; Conservative 2; Mismatches 108; Indels 4; Gaps 2;

QY 1 ATGCAATTTAGGCACTTTAGCTACAAATTTTAGCTCCCTGATTCGTTGGTGGCAACGAT 60
   |||||
Db 114 ATGAGAGAGTGACCTGCGAGGACCTCCCGAGCGCCACCATCGCTGTCACCTGGACCCG 173
   |||||

QY 61 GATGTTCTTACGCAAGAGTGAGACCGAGGCGCAAAATTTGAATCCCTCTTCAGAAACATATGAC 120
   |||||
Db 174 CGCGTGTCTGGACGCGCTGTGCGGCGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 233
   |||||

QY 121 AAGGACACCACTTCAGTATTTTAAAGAGCTTCAAGAGTCCCGGATPAAATCTTCAGCAAC 180
   |||||
Db 234 AAGGACATCACTTTTCAATATTTTAAAGAGCTTCAAGAGTCCCGGATPAAATCTTCAGCAAC 293
   |||||

QY 181 CCCTATCTCAGCGGATGCGAGTCCGCGCTGCACAGACCGAGTCCCTGGGAGAGAA 240
   |||||
Db 294 CCCTTCTCCGAGAGATGCGAGTCCGCGCTGCACAGACCGAGTCCCTGGGAGAGAA 353
   |||||

QY 241 ATGAAGTTGATTTTGTCTAGACTTTACACATAGGAAGTTTACACCTGGCTCCGCGCCAAAT 300
   |||||
Db 354 ATGAAGTTATATTTTGTCTAGACTTTACACATAGGAAGTTTACACCTGGCTCCGCGCCAAAT 412
   |||||

QY 301 CCGGACAAACAGTTCCTATCTCCCTCCGCGCTCTCTCCCTCCGCTGGCTGGGAAACAGTA 360
   |||||
Db 413 CCAGACAAACAGTTCCTATCTCTCCCTCCGCGCTCTCTCCGCTGGCTGGGAAACAGTG 472
   |||||

QY 361 GAAGATGCCACCCCGCTCATAAATTTAGATCTTTTATATGCCATCTCCAGCTGGGCGCA 420
   |||||
Db 473 GAAGATGCCACCCCGCTCATAAATTTAGATCTTTTATATGCCATCTCCAGCTGGGCGCA 532
   |||||

QY 421 GGAGAGAAGTATGAATGCAATGACGAGACAGCCCACTCCCACTGGTGGTGGTCCACGTG 480
   |||||
Db 533 GGGGAAAAGTATGAATTTGACGAGCGAGCTGACACCACTCCCACTGGTGGTGGTCCATGTA 592
   |||||

QY 481 TGTGAGTACCAACAGAGATGAGGAGAGAGAGAGAGATGAGAGATGAAGAGACCC 540
   |||||
Db 593 TGTGAGTATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649
   |||||

QY 541 AAGCCCAAAATATCAAGACACCGAGACCGAGTACACACCGATCCACCTTAGCTGA 597
   |||||
Db 650 AAGCCCAAAATATCAAGACACCGAGACCGAGTACACACCGATCCACCTTAGCTGA 706
   |||||

RESULT 9
US-09-954-456-497
; Sequence 497, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
```

```
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 497
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-497

Query Match      35.9%; Score 214.2; DB 10; Length 3184;
Best Local Similarity 66.1%; Pred. No. 9.8e-57;
Matches 331; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

QY 21 CTACAAATTTAGCTCCCTGATTCGTTGGTGGCAACGATGATGTTCTTCAGCGAAAGTGA 80
   |||||
Db 210 CTGTGATGTTTCCACTCTGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 269
   |||||

QY 81 GACCAAGGGGCAAAATTTGAATCCCTCTTCAGAAACATATGACAGGACACCACTTCCAGTA 140
   |||||
Db 270 GGTTAAGGAAAATTTGGGGGACTGTTTCGGACTTATGATGATGATGATGATGATGATGATGATG 329
   |||||

QY 141 TTTTAAGAGCTTCAAACGTTGTCGGATAAACTTCAGCAACCCCTTTATCTGCAGCCGATGC 200
   |||||
Db 330 ATTTAAGAGTTTTCAGACGTTGTCGGATAAACTTCAGCAATCTTAAATCTGCAGCCGATGC 389
   |||||

QY 201 CAGGCTCGCGCTGCACAGACCGAGTTCCTGGGAGAGAAATGAAGTTGTATTTTGTCTCA 260
   |||||
Db 390 TAGGATAGAGCTTCATGAACCCCAATTCAGAGGGGAAAAAATTAAGCTCTACTTTGCACA 449
   |||||

QY 261 GACT-----TTACACATAGGAAGTTTACACCTGGCTCCGCGCCCAATCCCGACAA 308
   |||||
Db 450 GGTTCAGACTCCAGACAGAGATGGAGCAAACTGCACTTGGCTCCACCCAGCTTGCCTCA 509
   |||||

QY 309 ACAGTTTCTCATCTCCCTCCGCGCTCTCTCTCCCTGGTGGCTGGAAACAAAGTAGAAGATGC 368
   |||||
Db 510 ACAGTTTCTCATCTCCCTCCGCGCTCTCTCTCCCTGGTGGCTGGAAACAAAGTAGAAGATGC 569
   |||||

QY 369 CACCCCGCTCATTAATTAGCATCTTTTATATGCCATCTTCAAGCTGGGCGCCAGAGAGAA 428
   |||||
Db 570 CACGCGAGTCTCAACTATGATGCTCTCTATGCTGTGGCCAAACTAGGACCCAGGAGAGAA 629
   |||||

QY 429 GTATGAATGCAATGACGAGCAGACAGCCCACTCCCACTGGTGGTGGTGGTGGTGGTGGTGGTGG 488
   |||||
Db 630 GTATGAGCTCCATGACGAGGACTGAGTCCACCCCAAGTGTCTGTCGTCGTCGTCGTCGTCGTCG 589
   |||||

QY 489 TCACCAAGAGATGAGGAGGA 509
   |||||
Db 690 TCACATAGAGAGAGAGAGGA 710
   |||||

RESULT 10
US-09-782-953-17
```

```

Sequence 17 Application US/09782953
Patent No. US20020150953A1
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 3184
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)..(780)
US-09-782-953-17

```

```

Query Match 331: 35.9%; Score 214.2; DB 10; Length 3184;
Best Local Similarity 66.1%; Pred. No. 9.9e-57;
Matches 331; Conservative 0; Mismatches 156; Indels 12; Gaps

Oy 21 CTACAAATTTAGCTCCCTGATGCTCTTGTTGTGGCAAAACGATGATGCTTTCAGCGAAAGTGA 80
Db 210 GTGTATGTTTTCACCTGTGTTGCTCTGTGTGTGTGATGTGCGAGGCTTTTACCAATCAGGA 269
Oy 81 GACCAAGGCGCAAAATTTGSAATCCCTCTTCAGACATATGACAAAGAGACACACCTTCCAGTA 140
Db 270 GATTAAAGAAAAATTTGGGGGACTGTTTCGACTTATGATGACTGTGTGACCTTCCAGCT 329
Oy 141 TTTTAAAGCTTCAAAACGTGTCCGATATMACTTCAGCAACCCCTTATCTGCAGCCGATGC 200
Db 330 ATTTAAAGATTTCAGACCGTGCTCGTATMAACTTCAGCAATCTTAATCTGCAGCCGAGC 389
Oy 201 CAGCGTCGCGCTGCACAAAGACCGAGTTCTGCGGGAGAAATGAAAGTGTATTTTGTCTCA 260
Db 390 TAGGATAGAGCTTCTAGAAACCCCAATTCAGAGGAAAAAATTTAAAGCTCTTACTTTGCACA 449
Oy 261 GACT-----TTACACATAGGAAGTTACACACCTGCTCCGCCAATCCCGAATA 308
Db 450 GGTTCAGACTCCAGAGACAGATGGAGACAAATGCACTTGSGTCCACCCCGACTGCCAA 509
Oy 309 ACAGTTCTCATCTCCCTCGGAGCTTCTCTCCGTTGGCTGGAAAACAAGTGAAGATGC 368
Db 510 ACAGTTTCTCATCTCGAGCCCTTCTCCGCCACCTTATAGCTGGCAGGCCATCAACGATGC 569
Oy 369 CACCCCGCTCAATAATTTAGACTTTTAAATGCCATCTCCAAAGCTGGGGCCAGGAGATA 428
Db 570 CACGCGAGCTCTCAACTATGACCTCTCTATCTGTGGCAAACTAAGACAGAGAGAGA 629
Oy 429 GATGTAAGCTGATGAGAGCAGACAGACCCCACTCCCAAGTGTGTGTGTCCAAGTGTGTAGAG 488
Db 630 GATATGAGCTCCATGACAGGAGACTGATGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 689
Oy 489 TGACCAAGAGAAATGAGAGAGA 509
Db 690 TGACATAGAGAAAGAGAGA 710

RESULT 11
US-09-782-953-8
; Sequence 8, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; BINDING OF INVENTION: CALCIUM-INDUCING INTERACTING PROTEIN (MCIIP)

```

```

1 FILE REFERENCE: UTSD:674P21
2 CURRENT APPLICATION NUMBER: US/09/782,953
3 CURRENT FILING DATE: 2001-02-13
4 PRIORITY APPLICATION NUMBER: 60/216,601
5 PRIOR FILING DATE: 2000-07-07
6 NUMBER OF SEQ ID NOS: 27
7 SOFTWARE: PatentIn Ver. 2.1
8 SEQ ID NO 8
9 LENGTH: 594
10
11 TYPE: DNA
12 ORGANISM: Mus musculus
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: (1)..(591)
16
17 US-09-782-953-8

```

Query Match	33.7%	Score	201.4	DB	10	Length	594
Best Local Similarity	62.8%	Pred.	No. 3, 76-53				
Matches	360	Conservative	0	Mismatches	186	Indels	27
				Gaps			2

  

QY	21	CTACAAATTTAGCTCCCTGATCTTGCTTGCTGGCAACGATGATGCTCTTCACGCAAAAGTGA	80
DB	21	CTGATGATTTTCACTGCTGCTGCTGTGTGGATGAGAGGCTTTTACCAATCAGGA	80
QY	81	GACCAAGGCCAAATTTGAAATCCCTCTTCAGAAACATATACAAAGGACACACCCTTCAGTA	140
DB	81	GATTAAAGAAAATTCGAGGGACGTGTTCCGAGCCATATATGAAATGTGTAGCTTCACCT	140
QY	141	TTTTAAGAGCTTCAACCGTGTCCGGATTAAGCTTCAGCAACCCCTTATCTGAGCCGATGC	200
DB	141	GTTTAAAGATTTCCGACGGGTTCCAAATTTTACGCCATCCCAAAATCTGACGCCGCTGC	200
QY	201	CAGCTCGGCTGCAAGAACCGAGTTCTCTGGGGAGGAAATGAATGTTATTTTGTCTCA	260
DB	201	CCGGATAGAGCTTCACTGAGACTCAAGTTCAAGAGGGAAGGCTAAAACTTACTTGCCCA	260
QY	261	G-----ACTTACATAGGAAGTTCAACCTGGGCTCCGCCAATCCGACAA	308
DB	261	GATCCAGACCCCGAGAGAGATGAGACAAACGTGCAATTTGGACCTCCACAGCTTGCCA	320
QY	309	ACAGTTCTTCATCTCCCTCCGGCCTCTCTCTCCCTGGTGGTGAACAAATGAAAGATGC	368
DB	321	ACAGTTCTTCATCTCAACCCCTTATCTCTCTCTGTGGTGGAAAGCCTATCAGGATGC	380
QY	369	CACCCCGCTCATTAATTACGATCTTTTATATGCCATCTTCCAAAGCTGGGGCCAGAGAGAA	428
DB	381	CACACACAGTCCCAACTATGACCTTCTTTATGCTGTGGCCAAACTAGGACCCGGAGAGAA	440
QY	429	GTAATGACTGCATGGAGAGACAGACCCCACTCCAGTGTGTTGTTGCCAAGTGTGAGAG	488
DB	441	ATTATAGCTGCAGCTGGAACCTGAACTCTACACCGACGTCGTGTGTGCATGTGTGTGACAG	500
QY	489	TGACCAAGAAATGAGAGGAGAAAGAGAAAGATGAGAGAAATGAAGACCCCAAGCCCAA	548
DB	501	CGACATGGAGGAGAGAA-----GGACCCAAAGACTTCCCCCAAGCCAAA	545
QY	549	AATCATCCAGACGAGAGACCGGATTAACACC	581
DB	546	AATCATTCAGACCCGCGCTCCGGGCTGTGCACCC	578

```

RESULT 12
US-09-782-953-23
/ Sequence 23 Application US/09782953
/ Patent No. US20020150953A1
/ GENERAL INFORMATION:
/ APPLICANT: WILLIAMS, R. SANDERS
/ APPLICANT: ROTHERMEL, BEVERLY
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
/ TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
/ FILE REFERENCE: USDS.674P21
/ CURRENT APPLICATION NUMBER: US/09/782,953
/ CURRENT FILING DATE: 2001-02-13

```



; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(637)  
US-09-782-953-23

Query Match 32.0%; Score 191.2; DB 10; Length 720;  
Best Local Similarity 63.2%; Pred. No. 6.4e-50;  
Matches 320; Conservative 0; Mismatches 168; Indels 18; Gaps 1;  
QY 4 GATTTAGGACTTTAGCTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACGATGAT 63  
Db 56 GAGATGATGATTTAGTGATCTGCCTACCTCACTTTTTCCTTTCGAGCGTCCATGAAGCA 115  
QY 64 GTCTTCAGCAAGTAGAGACCGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAG 123  
Db 116 GTGTTTGAGGCACGAGACAGAGAAAGATTTGAAGCACTCTTCACCATCTATGATGAC 175  
QY 124 GACACCACTTCCAGTATTTAGAGCTTCAAGAGTGTCCGGATAAATCTTCAGCAACCCC 183  
Db 176 CAGGTTACTTTTCAGCTGTTTAAAGCTTTAGAGAGTCAAGATAAATTTTCAGCAACCT 235  
QY 184 TTATCTGACGCGATGCGGCTGCACAGACCGAGTTCTGGGGAAGGAATG 243  
Db 236 GAAGCGGACAGAGCGCAATAGAACTCCAGAAACAGACTTCAATGGGCAAGCTA 295  
QY 244 AAGTGTATTTGCTCAGACTTTTACACATAGGAAGTTTCAACCTGGCTCGGCCCAATCCC 303  
Db 296 AAGCTATATTTGCAAG-----TCTATCTCTCGCGCCCGACCT 337  
QY 304 GACAAACAGTTCTCATCTCCCTCCGGCTCTCTCCCGCTCTCTCCCGTGGTGGAAACAGTAGAA 363  
Db 338 GTCAGAGAGTTCTCATCTCCCTCCAGCTCTCCCGAGTGGGTGGAGCAGAGCGAA 397  
QY 364 GATGCCACCCCGTCATAAATTTAGATCTTTTATATGCCATCTCCAGCTGGGCGCCAGGA 423  
Db 398 GATGCGAGCTCTTATATATTTAGTTACTCTGTGCTGTTTCCAAATGGGACACGGA 457  
QY 424 GAGAAGTATGAATGCTATGATGAGCAGACCGCCCACTCCCAAGTGTGTGTGTCACAGTGTGT 483  
Db 458 GAGAAATATGAATCTTACGCGGGAACAGAGTCCACACCCAGCGTGTGTCTGTCTGT 517  
QY 484 GAGAGTGACCAAGAGATGAGGAGGA 509  
Db 518 GAAAGTGAACCTGAAGAGGAAGA 543

RESULT 13  
US-09-782-953-20  
; Sequence 20, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: CALCIUM NEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: US/09/782,953  
; CURRENT APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; LENGTH: 828  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (23)..(745)  
US-09-782-953-20

Query Match 32.0%; Score 191.2; DB 10; Length 828;  
Best Local Similarity 62.4%; Pred. No. 6.9e-50;  
Matches 323; Conservative 0; Mismatches 183; Indels 12; Gaps 1;

QY 4 GATTTAGGACTTTAGCTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACGATGAT 63  
Db 134 GAGATGATGATTTAGTGATCTGCCTACCTCACTTTTTCCTTTCGAGCGTCCATGAAGCA 193  
QY 64 GTCTTCAGCGAAAGTAGAGACCGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAG 123  
Db 194 GTGTTTGAGGCACGAGACAGAGAAAGATTTGAAGCACTCTTCACCATCTATGATGAC 253  
QY 124 GACACCACTTCCAGTATTTTGAAGCTTCAAACTGTCCGGATAAATCTTCAGCAACCCC 183  
Db 254 CAGGTTACTTTTCAGCTGTTTAAAGCTTTAGAGAGTCAAGATAAATTTTCAGCAACCT 313  
QY 184 TTATCTGACGCGATGCGGCTGCACAGACCGAGTTCTTCGCGGAAGGAATG 243  
Db 314 GAAGCGGACAGAGCGGCAATAGAACTCCACGAAACAGACTTCAATGGGCGAGAGCTA 373  
QY 244 AAGTGTATTTTTCCTCAG-----ACTTTACACATAGGAAGTTTCACACCTGGCT 291  
Db 374 AAGCTATATTTTGCACAGTGTCCGCGAAGTCCGCGACAGTCTCTATCTCTG 433  
QY 292 CGGCCCAATCCCGACAAACAGTTCTCATCTCCCTCCGGCTCTCTCTCCCGTGGCTGG 351  
Db 434 CGGCCCGACCTGTCAAGCAGTTCTCATCTCCCTCCAGCTCTCTCCCGAGTGG 493  
QY 352 AAACAAGTAGAGATGCCACCCCGCTCATATAATTTAGATCTTTTATATGCCATCTCCAG 411  
Db 494 AAGCAGAGGAAGTGCAGTGTCTTTATAAATTTAGATTTTACTGTGTCTGTCTCCAA 553  
QY 412 CTGGGCGCAGGAGAGAGTATGAATGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAG 509  
Db 554 TTGGGACAGGAGAGAGATATGAATTTACGCGGGAACAGAGTCCACCCAGCGTGGT 613  
QY 472 GTCCACGTTGTGAGAGTGACCAAGAGATGAGGAGGA 509  
Db 614 GTTCATGTCTGTGAAGTGAAGTGAAGAGGAAGA 651

RESULT 14  
US-09-864-761-10388/c  
; Sequence 10388, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aconica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 10388
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000054.1
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67
US-09-864-761-10388
```

```

Query Match      24.2%; Score 144.6; DB 10; Length 412;
Best Local Similarity 89.1%; Pred. No. 1.7e-35;
Matches 156; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 87 GGGCAATTGATCCCTCTTGAACATATGACAGACACACCTTCAGATTATTTAA 146
DB 411 GGGCAATTGATCCCTCTTGAACATATGACAGACATACCTTCAGATTATTTAA 352
QY 147 GAGCTTCAACAGTGTCCGATTAACCTTATCTGACGCCGATGCCAGGCT 206
DB 351 GAGCTTCAACAGTGTCCGATTAACCTTATCTGACGCCGATGCCAGGCT 292
QY 207 GCGGCTGACAGACCGAGTTCTTGGGAAAGAAATGAAGTTGATTTGCTCAG 261
DB 291 CCAGCTGACATAGACTGAGTTCTTGGGAAAGAAATGAAGTTGATTTGCTCAG 237

RESULT 15
US-09-864-761-2064/c
Sequence 2064, Application US/09864761
Patent No. US2002004873A1
GENERAL INFORMATION:
APPLICANT: Penn. Sharon G.
APPLICANT: Rank. David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomic-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
```

```

PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 2064
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000122.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
US-09-864-761-2064
```

```

Query Match      24.2%; Score 144.6; DB 10; Length 446;
Best Local Similarity 89.1%; Pred. No. 1.7e-35;
Matches 156; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 87 GGGCAATTGATCCCTCTTGAACATATGACAGACACACCTTCAGATTATTTAA 146
DB 445 GGGCAATTGATCCCTCTTGAACATATGACAGACATACCTTCAGATTATTTAA 386
QY 147 GAGCTTCAACAGTGTCCGATTAACCTTATCTGACGCCGATGCCAGGCT 206
DB 385 GAGCTTCAACAGTGTCCGATTAACCTTATCTGACGCCGATGCCAGGCT 326
QY 207 GCGGCTGACAGACCGAGTTCTTGGGAAAGAAATGAAGTTGATTTGCTCAG 261
DB 325 CCAGCTGACATAGACTGAGTTCTTGGGAAAGAAATGAAGTTGATTTGCTCAG 271

Search completed: December 14, 2002, 20:01:42
```

Sun Dec 15 08:38:26 2002

Job time: 24.3119 secs

us-09-782-953-5.rnpb

---

Page 9

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:09:21 ; Search time 52.489 Seconds  
(without alignments)  
2732.361 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSWDCVSTLVACWVDV.....SPKPKIIQTRRPLGPSVSN 197

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Xgapop 10.0	Xgapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DBV=xlh  
-Q=/cg2\_1/USPTO.spool/US09782953/runat\_11122002\_114431\_17431/app\_query\_fasta\_1.1173  
-DB=Pending Patents NA New -QPMT=fastap -SUFFIX=rnnp -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09782953 -CGCN\_1\_1\_70/runat\_11122002\_114431\_17431 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : Pending Patents NA New :

1:	/cg2_5/prodata/1/pna/PCT_NEW_COMB.seq:
2:	/cg2_6/prodata/1/pna/US06_NEW_COMB.seq:
3:	/cg2_6/prodata/1/pna/US07_NEW_COMB.seq:
4:	/cg2_6/prodata/1/pna/US08_NEW_COMB.seq:
5:	/cg2_6/prodata/1/pna/US09_NEW_COMB.seq:
6:	/cg2_6/prodata/1/pna/US10_NEW_COMB.seq:
7:	/cg2_6/prodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	950	91.3	2411	5	US-09-724-676-15329
2	950	91.3	2411	5	US-09-724-676-15329
3	910	87.4	1021	6	US-09-724-676A-15329
4	905	86.9	3159	6	US-10-290-438-1
5	641	61.6	1931	5	US-09-724-676-12748
6	641	61.6	1931	5	US-09-724-676A-12748
7	641	61.6	2050	5	US-09-724-676-12756
8	641	61.6	2050	5	US-09-724-676A-12756
9	641	61.6	2297	5	US-09-724-676A-12747
10	641	61.6	2297	5	US-09-724-676A-12747
11	628	60.3	2348	6	US-10-240-965-60

12	626.5	60.2	1893	5	US-09-724-676-12741	Sequence 12741, A
13	626.5	60.2	1893	5	US-09-724-676A-12741	Sequence 12741, A
14	600	57.6	1839	5	US-09-724-676-12743	Sequence 12743, A
15	600	57.6	1839	5	US-09-724-676A-12743	Sequence 12743, A
16	598	57.4	1876	5	US-09-724-676-12745	Sequence 12745, A
17	598	57.4	1876	5	US-09-724-676A-12745	Sequence 12745, A
18	595	57.2	615	6	US-10-290-438-8	Sequence 8, Appli
19	482.5	46.3	911	5	US-09-724-676-12749	Sequence 12749, A
20	482.5	46.3	911	5	US-09-724-676A-12749	Sequence 12749, A
21	470	45.1	1094	5	US-09-724-676-12750	Sequence 12750, A
22	470	45.1	1094	5	US-09-724-676A-12750	Sequence 12750, A
23	470	45.1	1213	5	US-09-724-676-12740	Sequence 12740, A
24	470	45.1	1213	5	US-09-724-676A-12740	Sequence 12740, A
25	455.5	43.8	1056	5	US-09-724-676-12742	Sequence 12742, A
26	455.5	43.8	1056	5	US-09-724-676A-12742	Sequence 12742, A
27	429	41.2	1002	5	US-09-724-676-12744	Sequence 12744, A
28	429	41.2	1002	5	US-09-724-676A-12744	Sequence 12744, A
29	427	41.0	1039	5	US-09-724-676-12746	Sequence 12746, A
30	427	41.0	1039	5	US-09-724-676A-12746	Sequence 12746, A
31	323	31.0	442	5	US-09-513-999C-1772	Sequence 1772, Ap
32	300	28.8	1577	5	US-09-724-676-12755	Sequence 12755, A
33	300	28.8	1577	5	US-09-724-676A-12755	Sequence 12755, A
34	300	28.8	1943	5	US-09-724-676-12752	Sequence 12752, A
35	300	28.8	1943	5	US-09-724-676A-12752	Sequence 12752, A
36	294	28.2	531	6	US-10-290-438-6	Sequence 6, Appli
37	247.5	23.8	465	5	US-09-620-607B-1320	Sequence 1320, Ap
38	223	21.4	446	6	US-10-203-138A-8201	Sequence 8201, Ap
39	169	16.2	486	6	US-10-203-138A-827	Sequence 827, App
40	145	13.9	297	5	US-09-513-999C-8585	Sequence 8585, Ap
41	141.5	13.6	557	5	US-09-724-676-12753	Sequence 12753, A
42	141.5	13.6	557	5	US-09-724-676A-12753	Sequence 12753, A
43	131	12.6	123	6	US-10-203-138A-5980	Sequence 5980, Ap
44	129	12.4	740	5	US-09-724-676-12754	Sequence 12754, A
45	129	12.4	740	5	US-09-724-676A-12754	Sequence 12754, A

#### ALIGNMENTS

RESULT 1  
US-09-724-676-15329  
; Sequence 15329, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09724.676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 15329  
; LENGTH: 2411  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-15329

Alignment Scores:  
Pred. No.: 6.39e-87 Length: 2411  
Score: 950.00 Matches: 180  
Percent Similarity: 98.40% Conservative: 5  
Best Local Similarity: 95.74% Mismatches: 3  
Query Match: 91.26% Indels: 0  
DB: 5 Gaps: 0

US-09-782-953-9 (1-197) x US-09-724-676-15329 (1-2411)

Qy	1	MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal	20
Db	190	ATGCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGGTTCCTGTGGTGATGTC	249
Qy	21	GluValPheThrAsnGlnGluValIleGluValPheGluGlyLeuPheArgThrTyrAsp	40
Db	250	GAGGTCCTTTACCAATCAGGAGGTTAGGAAAAATTTGGGGACTGTTTCGACTTATGAT	309

```
Qy 41 GUCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
    : : : : :
Db 310 GACTGTGACGCTTCCAGCTATTAAAGTTTCACAGCGTGCCGTATAAACTTCAGCAAT 369
Qy 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisArgIleThrGlnPheArgGlyLys 80
    CCTAAATCTGACGCCGAGCTAGATAGAGCTTCATGTAACCAATTCAGAGGAAAAAA 429
Db 370
Qy 81 LeuLysLeuThrPheAlaGlnValGlnThrProGlnUtrrAspGlyAspLysLeuHisLeu 100
    : : : : :
Db 430 TTAAGCTCTTACTTGGACAGGTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTTG 489
Qy 101 AlaProGlnInProAlaLysGlnPheLeuLysSerProPheSerProProValGly 120
    : : : : :
Db 490 GCTCCACCCCACTGCGCAACAGTTTCTCATCTCGCCCTTCCTCCCACTGTTGGC 549
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTrpAspLeuLeuTrpAlaValAla 140
    : : : : :
Db 550 TGGCAGCCCATCAAGATGCCAGCGCAGTCTCTCACTATGACTCTCTATGCTGGCC 609
Qy 141 LysLeuGlyProGlyGlyLysTrpGluLeuHisAlaGlyThrGlnSerThrProSerVal 160
    : : : : :
Db 610 AAACCTAGGACGAGAGAGAAATGATGACTTCATGCGAGACTGATGCCCAAGTGC 669
Qy 161 ValValHisValLysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
    : : : : :
Db 670 GTGCTGACAGTGTGCGACAGTGCATAGACAGAAAGAACCAAGACTTCCCAAG 729
Qy 181 ProLysIleIleGlnThrArgArg 188
    : : : : :
Db 730 CCAAAATCATCCAGGCTGATAGA 753
```

## RESULT 2

```
US-09-724-676A-15329
; Sequence 15329, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15329
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-724-676A-15329
```

## Alignment Scores:

Pred. No.:	6,39e-87	Length:	2411
Score:	950.00	Matches:	180
Percent Similarity:	98.40%	Conservative:	5
Best Local Similarity:	95.74%	Mismatches:	3
Query Match:	91.26%	Indels:	0
DB:	5	Gaps:	0

US-09-782-953-9 (1-197) x US-09-724-676A-15329 (1-2411)

```
Qy 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20
    : : : : :
Db 190 ATGCGACCCCTTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 249
Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGlnGluLysPheArgThrTyAsp 40
    : : : : :
Db 250 GAGGTCTTTACCAATCAGAGGTTAAAGAAATTTGGGGACGTGTTCCGACTATGAT 309
    : : : : :
Qy 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
    : : : : :
Db 310 GACTGTGACGCTTCCAGCTATTAAAGTTTCACAGCGTGCCGTATAAACTTCAGCAAT 369
    : : : : :
Qy 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisArgIleThrGlnPheArgGlyLys 80
    : : : : :
```

```
Db 370 CCTAAATCTGACGCCGAGCTAGATAGACTTCATGTAACCAATTCAGAGGAAAAAA 429
    : : : : :
Qy 81 LeuLysLeuThrPheAlaGlnValGlnThrProGlnUtrrAspGlyAspLysLeuHisLeu 100
    : : : : :
Db 430 TTAAGCTCTTACTTGGACAGGTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTTG 489
Qy 101 AlaProGlnInProAlaLysGlnPheLeuLysSerProPheSerProProValGly 120
    : : : : :
Db 490 GCTCCACCCCACTGCGCAACAGTTTCTCATCTCGCCCTTCCTCCCACTGTTGGC 549
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTrpAspLeuLeuTrpAlaValAla 140
    : : : : :
Db 550 TGGCAGCCCATCAAGATGCCAGCGCAGTCTCTCACTATGACTCTCTATGCTGGCC 609
Qy 141 LysLeuGlyProGlyGlyLysTrpGluLeuHisAlaGlyThrGlnSerThrProSerVal 160
    : : : : :
Db 610 AAACCTAGGACGAGAGAGAAATGATGACTTCATGCGAGACTGATGCCCAAGTGC 669
Qy 161 ValValHisValLysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
    : : : : :
Db 670 GTGCTGACAGTGTGCGACAGTGCATAGACAGAAAGAACCAAGACTTCCCAAG 729
Qy 181 ProLysIleIleGlnThrArgArg 188
    : : : : :
Db 730 CCAAAATCATCCAGGCTGATAGA 753
```

## RESULT 3

```
US-10-290-438-3/c
; Sequence 3, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 09/614,474
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 219435_Rn.1
; US-10-290-438-3
```

## Alignment Scores:

Pred. No.:	2,28e-83	Length:	1021
Score:	910.00 <td>Matches:</td> <td>192</td>	Matches:	192
Percent Similarity:	94.63% <td>Conservative:</td> <td>2</td>	Conservative:	2
Best Local Similarity:	93.66% <td>Mismatches:</td> <td>3</td>	Mismatches:	3
Query Match:	87.42% <td>Indels:</td> <td>8</td>	Indels:	8
DB:	6 <td>Gaps:</td> <td>0</td>	Gaps:	0

US-09-782-953-9 (1-197) x US-10-290-438-3 (1-1021)

```
Qy 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20
    : : : : :
Db 855 ATGCGACCCCTTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 796
Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGlnGluLysPheArgThrTyAsp 40
    : : : : :
Db 795 GAGGTCTTTACCAATCAGAGGTTAAAGAAATTTGGGGACGTGTTCCGACTATGAT 736
    : : : : :
Qy 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
    : : : : :
Db 735 GATGTGACGCTTCCAGCTATTAAAGTTTCAGCGGCTCGAAATTAATTCAGCA 676
    : : : : :
```

QY 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHisGlu-ThrGlnPheArg-GlyL 79  
 Db 675 CCCCCAAGCTGCGAGCCGGTCCCGATAGAGCTTCATGTGGACCCAGTTTCATGAGGGA 616  
 QY 79 ysLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGly-AspLysLeu 98  
 Db 615 AGNAGCTGAAACTTACTTTTGCACAGGTCCAGACCCCGACAGACAGATGGTAGACAAACTG 556  
 QY 99 His-LeuAlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProPr 118  
 Db 555 CAGTTTGGCACCACCCACAGCTGCGCAACAGTTCCTCATCTCACCCCTTCATCTCTCC 496  
 QY 118 ovalGlyTrrLysProLysSerAspAlaThrProValLeuAsn-TyrAspLeuLeuTyrA 138  
 Db 495 GTTGGCTGGGAAGCTATCAGCGATGCCACACCGATCTCTCAACGTACGACCTCTTTATG 436  
 QY 138 laValAlaLysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrP 158  
 Db 435 CCGTGGCCAACTAGGACAGAGAGAAATATGAGCTGCATGCGGAACTGAGTCTACAC 376  
 QY 158 roSerValValHisValCysAspSerAspMetGluGluGluAspProLysThrS 178  
 Db 375 CGAGCGTTGTGTCACGTGTGTGACAGCACTTGGAGGAGGAGGATCCAAAGACTT 316  
 QY 178 erProLysProLysAlleleGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 315 CCCCAGCCAAAATCATCCAGACCCCGGCTCCTGGCTGCCACCCCTCGGTGTCCAAAC 257

## RESULT 4

US-10-290-438-1  
 ; Sequence 1, Application US/10290438  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loring, Jeanne F.  
 ; APPLICANT: Tingley, Debra W.  
 ; APPLICANT: Edwards, Carla M.  
 ; APPLICANT: Streeter, David G.  
 ; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN  
 ; FILE REFERENCE: PC-0013-1CIP  
 ; CURRENT APPLICATION NUMBER: US/10/290,438  
 ; CURRENT FILING DATE: 2002-11-06  
 ; PRIOR APPLICATION NUMBER: 09/614,474  
 ; PRIOR FILING DATE: 2000-07-11  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 1  
 ; LENGTH: 3159  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No: 247500.5  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: 216, 3091, 3093, 3103  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-10-290-438-1

Alignment Scores:  
 Pred. No.: 3,26e-82 Length: 3159  
 Score: 905.00 Matches: 171  
 Percent Similarity: 95.72% Conservative: 8  
 Best Local Similarity: 91.44% Mismatches: 8  
 Query Match: 86.94% Indels: 0  
 DB: 6 Gaps: 0

US-09-782-953-9 (1-197) x US-10-290-438-1 (1-3159)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 389 AACTCGTTGTTCGGTGCATGTTCACCGTACGTGTTTGAAGGAGAGAGAGAGGAA 448  
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50

Db 449 AAATTTGAGGAGCTGTTTCGGACTTATGACTGTGTGACGTTCCAGCTATTATTAGAGT 508  
 QY 51 PheArgValArgIleAlaenPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
 Db 509 TTCCAGACGTGTCCGTATAAACTTCAGCAATCTCTAAATCTGCAGCCGAGCTAGGATAGAG 568  
 QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
 Db 569 CTTTCATGAACCCCAATTCAGAGGGAAAAAATAAGCTCTACTTTGCACAGGTTTCAGACT 628  
 QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110  
 Db 629 CCAGACAGATGAGAGACAACTGCACTTTGGCTCCACCCAGCCCTGCCAAACAGTTTCTC 688  
 QY 111 IleSerProProSerSerProProValGlyTrrLysProLysSerAspAlaThrProVal 130  
 Db 689 ATCTGCCCCCTTCTCTCCACCTGTGTGGTGGCAGCCCATCAACGATGCCACGCGAGTC 748  
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
 Db 749 CTCCACTATGACTCTCTCTATGCTGTGGCCAACTAGGACCCAGGAGAGATATGAGCTC 808  
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170  
 Db 809 CATGCGAGGACTGAGTCCACCCCAAGTGTCTGTGGTGGCAGCTGTGGCAGACATGACATAGAG 868  
 QY 171 GluGluGluAspProLysThrSerProLysProLysLysIleGlnThrArgArgProGly 190  
 Db 869 GAAGAAGAGGAGCCCAAGACTTCCCAAGCCAAAATCATCCAAACTCGGCGCTCTGCGC 928  
 QY 191 LeuProProSerValSerAsn 197  
 Db 929 CTGCCACCTCTCGTGTCCAAAC 949

## RESULT 5

US-09-724-676-12748  
 ; Sequence 12748, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12748  
 ; LENGTH: 1931  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (11)-(11)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (18)-(18)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (23)-(23)  
 ; OTHER INFORMATION: n is a,c,g, or t  
 US-09-724-676-12748

Alignment Scores:  
 Pred. No.: 8,64e-56 Length: 1931  
 Score: 641.00 Matches: 120  
 Percent Similarity: 79.68% Conservative: 29  
 Best Local Similarity: 64.17% Mismatches: 30  
 Query Match: 61.58% Indels: 8  
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676-12748 (1-1931)

```

QY 11 SerThrLeuValAlaCysValAlaSerValAlaGluValPheThrAsnGlnGluValGlu 30
Db 93 AGCTCCCTGATTCCTGCTGCGCAAAAGTATATCTTCAGGAAAGGAAACGAGGCC 152
QY 31 LysPheGlnGluLeuPheArgThrTyraSpGluCysValThrPheGlnLeuPheLysSer 50
Db 153 AAATTGAGTCCCTCTTGTAGACGTATGCAAGACATCACCTTTCAGTATTTTAAGAGC 212
QY 51 PheArgValAlaGlnLeuPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 213 TTCAACGAGTACAGATTAATTAATTCACGAAACCCCTTCTCCGACGAGATCCAGGCTCAG 272
QY 71 LeuHisGlnThrGlnPheArgGlyLysLysLeuLysLeuTyrrPheAlaGlnValGlnThr 90
Db 273 CTGCATTAAGACTGAGTTCTGCGAAAGAAATGATTAATTTTGTCTCAGACCTTACAC 332
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 333 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCAAGTTTCTG 380
QY 111 IleserProProSerSerProProValGlyTrrPlyProIleserAspAlaThrProVal 130
Db 381 ATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAAAGTGAAGATCCGACCCAGTC 440
QY 131 LeuAsnTyrrAspLeuLeuTyrrAlaValAlaLysLeuGlyProGlyGluLysTyrrGluLeu 150
Db 441 ATAAACTATGATCTCTTATATGCACTCCAGCTGGGGCCAGGGGAAAGATGATGATG 500
QY 151 HisAlaGlyThrGlnSerThrProSerValValHisValCysAspSerAspMetGln 170
Db 501 CACGCGACGATGACACACACTCCAGCGTGTGTCCATGATGATGATGATGATGATG 560
QY 171 GlnGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 561 AAGGAGGAAGAAGAAATGGAAGAATGAGAGACCTAAGCCAAATTTATCCGAGAC 620
QY 187 ArgArgProGlyLeuProPro 193
Db 621 AGGAGCGCGAGTACACGCCG 641

RESULT 6
US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12748

Alignment Scores:
Pred. No.: 8,64e-56 Length: 1931
Score: 641.00 Matches: 120
Percent Similarity: 79.68% Conservative: 29

```

```

Best Local Similarity: 64.17% Mismatches: 30
Query Match: 61.58% Indels: 8
DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676A-12748 (1-1931)
QY 11 SerThrLeuValAlaCysValAlaSerValAlaGluValPheThrAsnGlnGluValGlu 30
Db 93 AGCTCCCTGATTCCTGCTGCGCAAAAGTATATCTTCAGGAAAGGAAACGAGGCC 152
QY 31 LysPheGlnGluLeuPheArgThrTyraSpGluCysValThrPheGlnLeuPheLysSer 50
Db 153 AAATTGAGTCCCTCTTGTAGACGTATGCAAGACATCACCTTTCAGTATTTTAAGAGC 212
QY 51 PheArgValAlaGlnLeuPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 213 TTCAACGAGTACAGATTAATTAATTCACGAAACCCCTTCTCCGACGAGATCCAGGCTCAG 272
QY 71 LeuHisGlnThrGlnPheArgGlyLysLysLeuLysLeuTyrrPheAlaGlnValGlnThr 90
Db 273 CTGCATTAAGACTGAGTTCTGCGAAAGAAATGATTAATTTTGTCTCAGACCTTACAC 332
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 333 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCAAGTTTCTG 380
QY 111 IleserProProSerSerProProValGlyTrrPlyProIleserAspAlaThrProVal 130
Db 381 ATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAAAGTGAAGATCCGACCCAGTC 440
QY 131 LeuAsnTyrrAspLeuLeuTyrrAlaValAlaLysLeuGlyProGlyGluLysTyrrGluLeu 150
Db 441 ATAAACTATGATCTCTTATATGCACTCCAGCTGGGGCCAGGGGAAAGATGATGATG 500
QY 151 HisAlaGlyThrGlnSerThrProSerValValHisValCysAspSerAspMetGln 170
Db 501 CACGCGACGATGACACACACTCCAGCGTGTGTCCATGATGATGATGATGATGATG 560
QY 171 GlnGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 561 AAGGAGGAAGAAGAAATGGAAGAATGAGAGACCTAAGCCAAATTTATCCGAGAC 620
QY 187 ArgArgProGlyLeuProPro 193
Db 621 AGGAGCGCGAGTACACGCCG 641

RESULT 7
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756

Alignment Scores:
Pred. No.: 9.35e-56 Length: 2050

```



Score: 641.00 Matches: 120  
 Percent Similarity: 79.68% Conservativity: 29  
 Best Local Similarity: 64.17% Mismatches: 30  
 Query Match: 61.58% Indels: 8  
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676-12756 (1-2050)

```

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 212 AGCTCCCTGATGCTGTGTCGCAACAGTATATCTTCAGCGAAGTGAACACAGGGCC 271
Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGCAAGACATCACCTTTTCAGTATTTTAAGAGC 331
Qy 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 332 TTCAACAGAGTCAGATAAACTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyPheAlaGlnValGlnThr 90
Db 392 CTGCATAAGACTGAGTTCTGGGAAGCAATGAAGTTATATTTTCTCAGACCTTACAC 451
Qy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 452 ATAGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAGCAGTTTCTG 499
Qy 111 IleSerProProSerSerProProValGlyTTrpLysProIleSerAspAlaThrProVal 130
Db 500 ATCTCCCTCCCGCTCTCCGCACTGGGATGGNACAGTGAAGATGCCACCCAGTC 559
Qy 131 LeuSerThrAspLeuLeuTyAlaValAlaLysLeuGlyProGlyGluLysTyGluLeu 150
Db 560 ATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGAAAGTATGAATTG 619
Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
Db 620 CACGACGAGTACACCATCTCCAGCGTGGTGTGCATGTATGTGAGATGATCAAGAG 679
Qy 171 GluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 680 AAGGAGGAAGAGGAATGGAAGATGAGGAGACCTTAAGCCAAATTTATCCAGACC 739
Qy 187 ArgArgProGlyLeuProPro 193
Db 740 AGGAGCGCGAGTACACGCCG 760

```

# RESULT 8

```

US-09-724-676A-12756
; Sequence 12756, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12756

```

## Alignment Scores:

Pred. No.: 9,358-56 Length: 2050  
 Score: 641.00 Matches: 120  
 Percent Similarity: 79.68% Conservativity: 29  
 Best Local Similarity: 64.17% Mismatches: 30  
 Query Match: 61.58% Indels: 8  
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676A-12756 (1-2050)

```

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 212 AGCTCCCTGATGCTGTGTCGCAACAGTATATCTTCAGCGAAGTGAACACAGGGCC 271
Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGCAAGACATCACCTTTTCAGTATTTTAAGAGC 331
Qy 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 332 TTCAACAGAGTCAGATAAACTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyPheAlaGlnValGlnThr 90
Db 392 CTGCATAAGACTGAGTTCTGGGAAGCAATGAAGTTATATTTTCTCAGACCTTACAC 451
Qy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 452 ATAGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAGCAGTTTCTG 499
Qy 111 IleSerProProSerSerProProValGlyTTrpLysProIleSerAspAlaThrProVal 130
Db 500 ATCTCCCTCCCGCTCTCCGCACTGGGATGGNACAGTGAAGATGCCACCCAGTC 559
Qy 131 LeuSerThrAspLeuLeuTyAlaValAlaLysLeuGlyProGlyGluLysTyGluLeu 150
Db 560 ATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGAAAGTATGAATTG 619
Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
Db 620 CACGACGAGTACACCATCTCCAGCGTGGTGTGCATGTATGTGAGATGATCAAGAG 679
Qy 171 GluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 680 AAGGAGGAAGAGGAATGGAAGATGAGGAGACCTTAAGCCAAATTTATCCAGACC 739
Qy 187 ArgArgProGlyLeuProPro 193
Db 740 AGGAGCGCGAGTACACGCCG 760

```

## RESULT 9

```

US-09-724-676-12747
; Sequence 12747, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t

```



; APPLICANT: TAI, Julie  
 ; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
 ; FILE REFERENCE: PA-0025 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/240,965  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: 60/195,106  
 ; PRIOR FILING DATE: 2000-04-05  
 ; NUMBER OF SEQ ID NOS: 276  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 60  
 ; LENGTH: 2348  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No: 042176.5  
 US-10-240-965-60

Alignment Scores:  
 Pred. No.: 2,318-54 Length: 2348  
 Score: 628.00 Matches: 120  
 Percent Similarity: 79.26% Conservative: 29  
 Best Local Similarity: 53.83% Mismatches: 30  
 Query Watch: 60.33% Indels: 9  
 DB: 6 Gaps: 2

US-09-782-953-9 (1-197) x US-10-240-965-60 (1-2348)

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
 Db 181 AGCTCCCTGATTGCTGTGGCAACAGTCATATCTTCAGCGAAAGTGAACAGGGCC 240  
 Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50  
 Db 241 AAATTGAGTCCCTCTTTAGACGCTATGACAAGGACATCACCTTTCAGTATTTTAAGAGC 300  
 Qy 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaAArgIleGlu 70  
 Db 301 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 360  
 Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLysLeuTyPheAlaGlnValGlnThr 90  
 Db 361 CTGCATAAGAGTCAGTTCTTGGGAAGAAATGAAGTTATATTTGCTCAGACTTACAC 420  
 Qy 91 ProGluThrAspGlyAspLysLeuHisLeu-AlaProProGlnProAlaLysGlnPheLe 110  
 Db 421 ATAGGAAGCTCA-----CACCTGGGCTCCGCCAAATCCAGACAAGCAGTTCT 468  
 Qy 110 uIleSerProProSerSerProProValGlyTyrPheProIleSerAspAlaThrProVa 130  
 Db 469 GATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAACAAGTGGAGATGCCAGCCAGT 528  
 Qy 130 LeuAsnTyAspLeuLeuTyAlaValAlaLysLeuGlyProGlyGlyLysTyGluLe 150  
 Db 529 CATAAATATGATCTCTTATATGCCATCTCCAGCTGGGCCAGGGGAAAGTATGATAT 588  
 Qy 150 uHisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetG 170  
 Db 589 GCACGAGCGCCAGTGCACACCATCCAGCGTGTGTGCTCATGTATGAGAGTGCATCAAGA 648  
 Qy 170 uGluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnTh 186  
 Db 649 GAAGGAGGAGAGAGAGAAATGAAGAATGAGGAGACCTTAAGCCAAATATATCCAGAC 708  
 Qy 186 rArgArgProGlyLeuProPro 193  
 Db 709 CAGGAGCGCGAGTACAGCGC 730

RESULT 12

US-09-724-676-12741  
 ; Sequence 12741, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 12741  
 ; LENGTH: 1893  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-12741

Alignment Scores:  
 Pred. No.: 2,478-54 Length: 1893  
 Score: 626.50 Matches: 121  
 Percent Similarity: 74.63% Conservative: 29  
 Best Local Similarity: 60.20% Mismatches: 34  
 Query Watch: 60.18% Indels: 17  
 DB: 5 Gaps: 3

US-09-782-953-9 (1-197) x US-09-724-676-12741 (1-1893)

Qy 6 MetAspCysAspVal-----SerThrLeuValAlaCys 16  
 Db 13 ATTGACTGGAGATGGAGGAGGTGGACCTCGAGGACCTGCCAGCGCCACCATCGCTGT 72  
 Qy 17 ValValAspValGluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPhe 36  
 Db 73 CACCTGGACCCCGGCTGTCTGTGACGGCTGTGCCGGGCCAAATTTGAGTCCCTCTTT 132  
 Qy 37 ArgThrTyAspGluCysValThrPheGlnLeuPheLysSerPheArgArgValArgIle 56  
 Db 133 AGGAGCTATGACAAGGACATCACCTTTCAGTATTTTACAGGCTTCAACGAGTCAAGATA 192  
 Qy 57 AsnPheSerHisProLysSerAlaAlaArgAlaAArgIleGluLeuHisGluThrGlnPhe 76  
 Db 193 AACCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTT 252  
 Qy 77 ArgGlyLysLeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAsp 96  
 Db 253 CTGGGAAGGAATGAAGTTATATTTGCTCAGACTTACACATAGGAAGCTCA----- 306  
 Qy 97 LysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuLysSerProProSerSer 116  
 Db 307 -----CACCTGGCTCCGCCAAATCCAGACAAGCAGTTCTGATCTCCCTCCCGCTCT 360  
 Qy 117 ProProValGlyTyrPheProIleSerAspAlaThrProValLeuAsnTyAspLeuLeu 136  
 Db 361 CGCCAGTGGGATGGAACAAGTGGAGATGCCAGCCAGTCATAAACTATGATCTCTTA 420  
 Qy 137 TyrAlaValAlaLysLeuGlyProGlyGlyLysTyGluLeuHisAlaGlyThrGluSer 156  
 Db 421 TATGCCATCTCCAGCTGGGCCAGGGGAAAGTATGAATTCAGCGCAGCCAGTGCACAC 480  
 Qy 157 ThrProSerValValHisValCysAspSerAspMetGluGluGluGluAspProLys 176  
 Db 481 ACTCCAGCGTGGTGTCCATGTATGTGAGAGTGCATCAAGAGAGGAGAGAGAGAGAA 540  
 Qy 177 ThrSer-----ProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192  
 Db 541 ATGGAAGAATGAGGAGACCTTAAGCCAAATATATCCAGACCGAGCGCGGAGTACACG 600  
 Qy 193 Pro 193  
 Db 601 CCG 603

RESULT 13

US-09-724-676A-12741  
 ; Sequence 12741, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen



Score:	600.00	Matches:	115
Percent Similarity:	79.77%	Conservative:	23
Best Local Similarity:	66.47%	Mismatches:	27
Query Match:	57.64%	Indels:	8
DB:	5	Gaps:	2

US-09-782-953-9 (1-197) x US-09-724-676A-12743 (1-1839)

QY	25	ASnGlnGluValysGluIysPheGluGlyLeuPheArgThrTyrAspGluCysValThr	44
DB	43	AACAGAAATGTTGGGCCAAATTTGAGTCCCTCTTTAGGCGGTATGACAGGACATCACC	102
QY	45	PheGlnLeuPheIysSerPheArgArgValArgIleAsnPheSerHisProIysSerAla	64
DB	103	TTTCAGTATTTTAAAGAGCTTCAACACAGGTGAGATAAACTTCAGCAACCCCTTCGCGCA	162
QY	65	AlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLeuIysLeuValTyr	84
DB	163	GCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTCGGAAGGAATGAAGTTATAT	222
QY	85	PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProGln	104
DB	223	TTTGCTCAGACCTTACATAGGAAGCTCA-----CAGCTGGCTCGGCCAAAT	270
QY	105	ProAlaLysGlnPheLeuIleSerProProSerSerProProValGlyTrrIysProIle	124
DB	271	CCAGACAACGAGTTCTGATCTCCCTCCCGCTCTCGCCAGTGGGATGGAAACAAGTG	330
QY	125	SerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyPro	144
DB	331	GAAGATGCCACCCAGTCATAAACTATGATCTCTTTATATGCCATCTCCAAAGCTGGGGCA	390
QY	145	GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValHisVal	164
DB	391	GGGGAAGAAGTATGAATTGACCACGCGACTGACACCACTCCAGCGTGGTGGTCATGTA	450
QY	165	CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys	180
DB	451	TGTGAGAGTGATCAAGAGAGAAGGAGGAAGAGGAATGGAAGAGNATGAGGAGACCTAAG	510
QY	181	ProLysIleIleGlnThrArgArgProGlyLeuProPro	193
DB	511	CCAAAATTTATCCAGCCAGGAGGCCGAGTAGTACACGCCG	549

Search completed: December 15, 2002, 00:04:31  
Job time : 58.489 secs

**THIS PAGE BLANK (USPTO)**



QY 131 LNYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCDSMEEDPKTS---PKPKIIQT 186  
 Db 127 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCESQDEKEBEMERMRPKKIIQT 186  
 QY 187 RRPGLPP 193  
 Db 187 RRPGLPP 193

RESULT 9  
 US-09-782-953-24  
 / Sequence 24, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHERMEL, BEVERLY  
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 / FILE REFERENCE: UTSD:674P21  
 / CURRENT APPLICATION NUMBER: US/09/782,953  
 / PRIOR FILING DATE: 2001-02-13  
 / PRIOR APPLICATION NUMBER: 60/216,601  
 / NUMBER OF SEQ ID NOS: 27  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 24  
 / LENGTH: 212  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 US-09-782-953-24

Query Match 60.9%; Score 633.5; DB 10; Length 212;  
 Best Local Similarity 67.9%; Pred. No. 8.5e-55;  
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

QY 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRRVRINFSHPKSAARARIE 70  
 Db 28 TSLFACSVHEAVFEARQKERFEALFTIYDDQVTFQLFKSFRRVRINFSKPEAAARARIE 87  
 QY 71 LHETOPRGKKLKYFAOVQTPETDGDKLHLAPPOAKQFLISPPSSPPVGMKPSIDATPV 130  
 Db 86 LHETDFNGQKLYFAQ-----STLLPQPVKQFLISPPSPVGMKQSEDAAPV 137  
 QY 131 LNYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCDSMEEDPKTSPPKIIQTRPG 190  
 Db 138 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCESQDEKEBEMERMRPKKIIQT 196

QY 191 LPPSVSN 197  
 Db 197 PPTAALN 203

RESULT 10  
 US-09-782-953-25  
 / Sequence 25, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHERMEL, BEVERLY  
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 / FILE REFERENCE: UTSD:674P21  
 / CURRENT APPLICATION NUMBER: US/09/782,953  
 / PRIOR FILING DATE: 2001-02-13  
 / PRIOR APPLICATION NUMBER: 60/216,601  
 / NUMBER OF SEQ ID NOS: 27  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 25  
 / LENGTH: 212  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens

US-09-782-953-25

Query Match 60.9%; Score 633.5; DB 10; Length 212;  
 Best Local Similarity 67.9%; Pred. No. 8.5e-55;  
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

QY 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRRVRINFSHPKSAARARIE 70  
 Db 28 TSLFACSVHEAVFEARQKERFEALFTIYDDQVTFQLFKSFRRVRINFSKPEAAARARIE 87  
 QY 71 LHETOPRGKKLKYFAOVQTPETDGDKLHLAPPOAKQFLISPPSSPPVGMKPSIDATPV 130  
 Db 86 LHETDFNGQKLYFAQ-----STLLPQPVKQFLISPPSPVGMKQSEDAAPV 137  
 QY 131 LNYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCDSMEEDPKTSPPKIIQTRPG 190  
 Db 138 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCESQDEKEBEMERMRPKKIIQT 196  
 QY 191 LPPSVSN 197  
 Db 197 PPTAALN 203

RESULT 11  
 US-09-782-953-6  
 / Sequence 6, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHERMEL, BEVERLY  
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 / FILE REFERENCE: UTSD:674P21  
 / CURRENT APPLICATION NUMBER: US/09/782,953  
 / PRIOR FILING DATE: 2001-02-13  
 / PRIOR APPLICATION NUMBER: 60/216,601  
 / NUMBER OF SEQ ID NOS: 27  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 6  
 / LENGTH: 198  
 / TYPE: PRT  
 / ORGANISM: Mus musculus  
 US-09-782-953-6

Query Match 60.8%; Score 632.5; DB 10; Length 198;  
 Best Local Similarity 64.4%; Pred. No. 9.7e-55;  
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRRVRINFSHPKSAARARIE 70  
 Db 11 SLIACVANDVDFSESETRAFESLFRYYDKDTTFQYFKSFRRVRINFSNPLISADARLR 70  
 QY 71 LHETOPRGKKLKYFAOVQTPETDGDKLHLAPPOAKQFLISPPSSPPVGMKPSIDATPV 130  
 Db 71 LHETDFNGQKLYFAQTLHGSS---HLAPNPDQKFLISPPSPVGMKQSEDAAPV 126  
 QY 131 LNYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCDSMEEDPKTSPPKIIQTRPG 185  
 Db 127 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCESQDEKEBEMERMRPKKIIQT 186  
 QY 186 TRRGLPP 193  
 Db 187 TRRGLPP 194

RESULT 12  
 US-09-782-953-7  
 / Sequence 7, Application US/09782953  
 / Patent No. US20020150953A1  
 / GENERAL INFORMATION:  
 / APPLICANT: WILLIAMS, R. SANDERS  
 / APPLICANT: ROTHERMEL, BEVERLY



```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match          60.8%; Score 632.5; DB 10; Length 198;
Best Local Similarity 64.4%; Pred. No. 9.7e-55;
Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 STLAVCVDEVTNOEVEKFEGLRTYDECVTFQPKSFRVRINFSPKSAARARIE 70
DB 11 SSLIACVANDVSESESTRAKFESLFTYDKDTTFQYFKSFRVRINFSPKSAADARLR 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKQFLISPPSPPVGWKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPSPPVGWKQVEDATPV 126
QY 131 LNYDLLYAVAKLPGEGKELHAGTSTPSVVHVCDSDMEEDPK-----TSPKPKIIQ 185
DB 127 INVDLLYAIKSLGPGKELHAGTSTPSVVHVCDSDMEEDPK-----TSPKPKIIQ 186
QY 186 TRRGGLPP 193
DB 187 TRPEYTP 194

RESULT 13
US-09-782-953-15
; Sequence 15, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

Query Match          59.8%; Score 622; DB 10; Length 197;
Best Local Similarity 63.6%; Pred. No. 1e-53;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLAVCVDEVTNOEVEKFEGLRTYDECVTFQPKSFRVRINFSPKSAARARIE 70
DB 11 SATIACHLDPRVFDGLCRKAFESLFTYDKDITFQYFKSFRVRINFSPKSAADARLQ 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKQFLISPPSPPVGWKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPSPPVGWKQVEDATPV 126
QY 131 LNYDLLYAVAKLPGEGKELHAGTSTPSVVHVCDSDMEEDPKTS-----PKPKIIQ 186
DB 127 INVDLLYAIKSLGPGKELHAGTSTPSVVHVCDSDMEEDPKTS-----PKPKIIQ 186
QY 187 TRRGGLPP 193
DB 187 TRPEYTP 193

RESULT 15
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

Query Match          58.9%; Score 613.5; DB 10; Length 198;
Best Local Similarity 63.3%; Pred. No. 6.9e-53;
```

```
QY 187 TRRGGLPP 193
DB 187 TRPEYTP 193

RESULT 14
US-09-782-953-16
; Sequence 16, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match          59.8%; Score 622; DB 10; Length 197;
Best Local Similarity 63.6%; Pred. No. 1e-53;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLAVCVDEVTNOEVEKFEGLRTYDECVTFQPKSFRVRINFSPKSAARARIE 70
DB 11 SATIACHLDPRVFDGLCRKAFESLFTYDKDITFQYFKSFRVRINFSPKSAADARLQ 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKQFLISPPSPPVGWKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPSPPVGWKQVEDATPV 126
QY 131 LNYDLLYAVAKLPGEGKELHAGTSTPSVVHVCDSDMEEDPKTS-----PKPKIIQ 186
DB 127 INVDLLYAIKSLGPGKELHAGTSTPSVVHVCDSDMEEDPKTS-----PKPKIIQ 186
QY 187 TRRGGLPP 193
DB 187 TRPEYTP 193

RESULT 15
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

Query Match          58.9%; Score 613.5; DB 10; Length 198;
Best Local Similarity 63.3%; Pred. No. 6.9e-53;
```

Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

```

QY 11 STLVACVVDEVEFTNQEVKEKPEGLFRITYDECVTFQLFKSPFRVRINFSPKSAARARIE 70
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 11 SATIACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSNPLSADARLR 70
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 71 LHETQFRGKUKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVGWKKPISDATPV 130
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 71 LKTEFLGKEMKLYFAQTLHIGSS---HLAPPNPKQFLISPPASPVGWKQVEDATPV 126
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 131 LNYDLLVAVAKLGPGKVELHAGTESTPSVVVHVCSDMEEREDPK----TSPKXIIQ 185
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 127 INYDLLVAISKLGPGKVELHAATDPTBSVVVHVCSDQENEEBEMERMRKRPKXIIQ 186
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 186 TRRPGGLP 193
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 187 TRRPEYTP 194
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Search completed: December 11, 2002, 11:50:42  
 Job time : 7.31197 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 32.8887 Seconds  
(without alignments)  
2368.771 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MRAPSMDCVSTLVACVVDV.....SPAPKIIQTRRGLPPSVSN 197

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09782953/runat\_11122002\_114431\_17465/app.query.fasta\_1.1173  
-DB=Published Applications NA -OFMT=fastap -SUPFI=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09782953@cgn 1.1.38 @runat\_11122002\_114431\_17465  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1041	100.0	594	10	US-09-782-953-8 Sequence 8, Appli
2	1007	96.7	3184	10	US-09-954-456-497 Sequence 497, App
3	1007	96.7	3184	10	US-09-782-953-17 Sequence 17, Appli
4	658.5	63.3	828	10	US-09-782-953-20 Sequence 20, Appli

#### ALIGNMENTS

RESULT 1

US-09-782-953-8

; Sequence 8, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; FILE REFERENCE: US/09782,953

; CURRENT APPLICATION NUMBER: US/09782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 8

; LENGTH: 594

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(591)

US-09-782-953-8

Alignment Scores:

Pred. No.:

Score: 6.35e-107

Length: 1041.00

Matches: 594

197

Sequence 11, Appli  
Sequence 255, App  
Sequence 23, Appli  
Sequence 5, Appli  
Sequence 14, Appli  
Sequence 2, Appli  
Sequence 340, Ap  
Sequence 347, App  
Sequence 337, Ap  
Sequence 1, Appli  
Sequence 10388, A  
Sequence 2084, Ap  
Sequence 736, App  
Sequence 809, App  
Sequence 27019, A  
Sequence 17592, A  
Sequence 18808, A  
Sequence 1474, Ap  
Sequence 361, App  
Sequence 422, App  
Sequence 422, App  
Sequence 422, App  
Sequence 536, App  
Sequence 4005, Ap  
Sequence 302, App  
Sequence 522, App  
Sequence 3371, Ap  
Sequence 9, Appli  
Sequence 3, Appli  
Sequence 27196, A  
Sequence 10575, A  
Sequence 105, App  
Sequence 1740, Ap  
Sequence 13, Appli  
Sequence 5722, Ap  
Sequence 1, Appli  
Sequence 1691, Ap  
Sequence 1691, Ap  
Sequence 583, App  
Sequence 10910, A  
Sequence 1690, Ap  
Sequence 1690, Ap  
Sequence 1883, Ap

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-782-953-9 (1-197) x US-09-782-953-8 (1-594)

QY 1 MecProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20  
 Db 1 ATGCCAGCCCCCTAGCATGACTGTGATGTTCCACCTGGTGGCTGTGTGTGATG 60  
 QY 21 GluValPheThrAspGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40  
 Db 61 GAGGCTCTTACCAATCAGAGGTTAGGAAAATTCAGAGGACTGTTCCGAGCTTATGAT 120  
 QY 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60  
 Db 121 GATGTGTACGTTCCAGCTGTTTAAGACTTTCGACGGGTTGAAATATTTCAAGCCAT 180  
 QY 61 PolySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLys 80  
 Db 181 CCCAATCTGCAGCCCGTCCCGATAGAGCTTCATGAGACTCAGTTCAGAGGGAAG 240  
 QY 81 LeuValLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
 Db 241 CTAAAACTCTACTTCCCGCAGGTCCAGACCCCAAGACAGATGAGACAACTGCATTGG 300  
 QY 101 AlaProProGlnProAlaLysGlnPheLeuLysSerProSerSerProProValGly 120  
 Db 301 GACCTCCACACCGCTCCCAACAGTCTCTCATCTCACCCCTTCATCTCTCTGTGGC 360  
 QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140  
 Db 361 TGGAGCCCTATCAGCATGCCACACAGTCCCTCACTATGACTTCTTATCTGTGGCC 420  
 QY 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGlnSerThrProSerVal 160  
 Db 421 AAACCTAGGACCGAGAGAAATATGAGCTGCACGCTGGAAGCTGAGTCTACCGAGAGCTTC 480  
 QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180  
 Db 481 GTGGTGCACTGTGTGACAGCGACATGAGAGGAGAGAGAACCCAAAGACTTCCCCCAAG 540  
 QY 181 PolyIleIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 541 CCAAAAATCATTCAGACCCGGCGCTCCGGGCTTCCACCTCCGTGTCCAAAC 591

RESULT 2  
 US-09-954-456-497  
 Sequence 497, Application US/09954456  
 Patent No. US20020115057A1  
 GENERAL INFORMATION:  
 APPLICANT: Young, Paul  
 TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
 FILE REFERENCE: 689290-76  
 CURRENT APPLICATION NUMBER: US/09/954,456  
 PRIOR FILING DATE: 2001-09-18  
 PRIOR APPLICATION NUMBER: US/60/233,617  
 PRIOR FILING DATE: 2000-09-18  
 PRIOR APPLICATION NUMBER: US/60/234,052  
 PRIOR FILING DATE: 2000-09-20  
 PRIOR APPLICATION NUMBER: US/60/234,923  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US/60/235,134  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US/60/235,637  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235,638  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235,711  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,720

PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,840  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,863  
 PRIOR FILING DATE: 2000-09-27  
 NUMBER OF SEQ ID NOS: 2276  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 497  
 LENGTH: 3184  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-954-456-497

Alignment Scores:  
 Pred. No.: 3.77e-102 Length: 3184  
 Score: 1007.00 Matches: 190  
 Percent Similarity: 98.98% Conservative: 5  
 Best Local Similarity: 96.45% Mismatches: 2  
 Query Match: 96.73% Indels: 0  
 Gaps: 0

US-09-782-953-9 (1-197) x US-09-954-456-497 (1-3184)

QY 1 MecProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20  
 Db 190 ATGCCAGCCCCCTAGCATGACTGTGATGTTCCACCTGGTGGCTGTGTGTGATG 249  
 QY 21 GluValPheThrAspGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40  
 Db 250 GAGGCTCTTACCAATCAGAGGTTAGGAAAATTTGGGGAGCTGTTCCGACTTATGAT 309  
 QY 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60  
 Db 310 GACTGTGAGAGTTCACACTTTAAGAGTTTCAGACCTGTCCTGATTAACCTCAGCAAT 369  
 QY 61 PolySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLys 80  
 Db 370 CTAAATCTGACGCCCGAGCTGAGTATGAGCTTCTTGAACCCCAATTTAGAGGAAAATA 429  
 QY 81 LeuValLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
 Db 430 TTAAGCTCTACTTGTGACAGGTTGACAGCTCCAGAGACAGATGAGAACAACTGCACCTTG 489  
 QY 101 AlaProProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120  
 Db 490 GCTCACCCACGCGCCCAACAGTTTCTCATCTGCCCTTCTCCCACTGTGTAGC 549  
 QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140  
 Db 550 TGGCAGCCCATCAGATGCCACGCAATCTCTCAACTATGACTTCTTATGCTGTGGCC 609  
 QY 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGlnSerThrProSerVal 160  
 Db 610 AAACCTAGGACCGAGAGAAATATGAGCTGCACGCTGGAAGCTGAGTCTACCGAGAGCTTC 669  
 QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180  
 Db 670 GTGGTGCACTGTGTGACAGCGACATGAGAGGAGAGAGAACCCAAAGACTTCCCCCAAG 729  
 QY 181 PolyIleIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
 Db 730 CCAAAAATCATTCAAACTGGGCTCTGTGGCTCCGACCCCTCGTGTCCAAAC 780

RESULT 3  
 US-09-782-953-17  
 Sequence 17, Application US/09782953  
 Patent No. US20020150953A1  
 GENERAL INFORMATION:  
 APPLICANT: WILKINS, R. SANDERS  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 FILE REFERENCE: UTSD:674P21

; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 3184  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (205)..(780)  
US-09-782-953-17

Alignment Scores:  
Pred. No.: 3,77e-102 Length: 3184  
Score: 1007.00 Matches: 190  
Percent Similarity: 98.98% Conservatives: 5  
Best Local Similarity: 96.45% Mismatches: 2  
Query Match: 96.73% Indels: 0  
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x US-09-782-953-17 (1-3184)

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20  
DB 190 ATCCACGCCCTAGCAGTGTGATGTTTCCACTCTGCTGTGGTGGATGTC 249  
QY 21 GluValPheThrAngGlnGluValLysGluLysPheGluLysPheArgThrTyrAsp 40  
DB 250 GAGGTCCTTACCACATCAGGAGGTTAAGAAAAATTTGGGGGAGCTGTTTCGACATTATGAT 309  
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgLeuAsnPheSerHis 60  
DB 310 GACTGTGTGAGCTTCCAGCTATTTAAGAGTTTCAGAGCTGTCCGTATAAATTCAGCAAT 369  
QY 61 ProLysSerAlaAlaArgAlaArgileGluLeuHisGlnThrGlnPheArgGlyLysLys 80  
DB 370 CCTAAATCTCAGCCCGAGCTAGATAGAGCTTCATGAACCCCAATTCAGAGGGAATAA 429  
QY 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100  
DB 430 TTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGGAGCAAACTGCACCTTG 489  
QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProValGly 120  
DB 490 GCTCCACCCAGCTGCCAACAAGTTCTCATCTCGCCCTCTCTCCCTCCCTCTGTAGC 549  
QY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140  
DB 550 TGGCAGCCCATCAACGATGCCAGCCAGTCTCACTATGACCTCTCTATGCTGTGGCC 609  
QY 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160  
DB 610 AAATAGGACCCAGAGAGAGATGATGCTCCATGCCAGGAGTCCAGTCCACCCCAAGTGC 669  
QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180  
DB 670 GTCGTCACCTGTGCGACAGTGCATAGAGAGAGAGAGACCCCAAGACTTCCCCAAG 729  
QY 181 ProLysLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197  
DB 730 CCAAAATATCAAACTCGCGCTCTGGCTGCCACCTCTCGTGTCCAAC 780

RESULT 4

US-09-782-953-20  
; Sequence 20, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 828  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (23)..(745)  
US-09-782-953-20

Alignment Scores:  
Pred. No.: 2,63e-64 Length: 828  
Score: 658.50 Matches: 131  
Percent Similarity: 79.14% Conservatives: 17  
Best Local Similarity: 70.05% Mismatches: 38  
Query Match: 63.26% Indels: 1  
DB: 10 Gaps: 1

US-09-782-953-9 (1-197) x US-09-782-953-20 (1-828)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30  
DB 161 ACTCTATCTTTTGTCTTGACGCTCCATGAAGCAGTGTGTTGAGCAGCAGAGAGAGGAA 220  
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
DB 221 AGATTGAAGCACTCTTCCACCTATGATGACCAGGTTACTTTTCAGCTGTTTAAAGC 280  
QY 51 PheArgArgValArgileAsnPheSerHisProLysSerAlaAlaArgAlaArgileGlu 70  
DB 281 TTAGAAGAGTGCAGATAAATTTTCAGAAACCTGAAGCGCAGCAAGAGCGCGAATAGAA 340  
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90  
DB 341 CTCACGAAACAGACTTCAATGGCGCAGAGCTAAAGCTATATTTTCCACAGGTGCAGATG 400  
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110  
DB 401 TCCGGCGAAGTGGCGACAGTCTCTATCTCTGCGCCGCCAGCCTGTCAGCAGTTCCTC 460  
QY 111 IleSerProProSerSerProProValGlyTyrLysProLysSerAspAlaThrProVal 130  
DB 461 ATCTCCCTCCAGCTCTCCCGCAGTGGGTGGAGCAGCAGCAGATCGCATGCTGTT 520  
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
DB 521 ATAAATATGATTTACTCTGTCTGTTTCCAAATTTGGGACCCAGGAGAGAAATATGAATT 580  
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170  
DB 581 CACCGCGAACAGAGTTCGACACCCAGCGTGTGTTCATGTTCTGTGAAGTGAACCTGAA 640  
QY 171 GluGluGluAspProLysThrSerProLysProLysLeuLysLeuLysLeuLysLeuLys 190  
DB 641 GAGGAAGAGAGAGACAAA---AACCCCAACAGAAAAATTTGCCAGACAGAGCGCCCGAC 697  
QY 191 LeuProProSerValSerAsn 197  
DB 698 CCTCCGACCCGACGCTTGAAT 718

RESULT 5

US-09-782-953-11  
; Sequence 11, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS

```

; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)..(734)
US-09-782-953-11

```

```

Alignment Scores:
Pred. No.: 9,54e-62 Length: 2331
Score: 641.00 Matches: 120
Percent Similarity: 79.68% Conservative: 29
Best Local Similarity: 64.17% Mismatches: 30
Query Match: 61.58% Indels: 8
DB: Gaps: 2

```

```

US-09-782-953-9 (1-197) x US-09-782-953-11 (1-2331)
QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValIleGlu 30
DB 174 AGCTCCCGATTGCTGCTGTGGCAACAGTGAATCTTACGCAAAAGTAAACCGAGGCC 233
QY 31 LysPheGlnGluValLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
DB 234 AAATTGAGTCCCTCTTTAGACGATGACAAAGACATACCTTTCACTATTAAAGAC 293
QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgIleGlu 70
DB 294 TTCAAAACGAGTCAAGATTAATCTTCAAGCAACCTTCTCCGACAGATCCAGGCTCCAG 353
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90
DB 354 CTGCATTAAGACTGAGCTTCTGGAAAGAAATGAATGATTATTTGCTCAACCTTACAC 413
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
DB 414 ATAGGAAGCTCA-----CACCTGCTCCGCCCAATCCAGACAAAGCATTTCTG 461
QY 111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130
DB 469 ATCTCCCTCCCGCTCTCTCCGCAAGTGGATGGAACAAGTGAATGCGACCCAGTC 521
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
DB 522 ATAAACTATGATCTCTTATATGCCATCTCCAGCTGGGCCAGGGGAAAGATGATG 581
QY 151 HisAlaGlyThrGlnLysSerThrProSerValValHisValCysAspSerAspMetGlu 170
DB 582 CACGACGACACTGACACCACTCCACGCTGGTGTCTCATGTATGATGATGATCAAGAG 641
QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
DB 642 AAGGAGGAAGAAGGAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 701
QY 187 ArgArgProGlyLysProPro 193
DB 702 AGGAGGCCGAGATACACGCCG 722

```

RESULT 6  
 US-10-044-090-255  
 ; Sequence 255, Application US/10044090  
 ; Patent No. US20020137081A1

```

; GENERAL INFORMATION:
; APPLICANT: Olga Sandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO: 255
; LENGTH: 2355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incycle ID No. US20020137081A1 5511889CB1
US-10-044-090-255

```

```

Alignment Scores:
Pred. No.: 9.68e-62 Length: 2355
Score: 641.00 Matches: 120
Percent Similarity: 79.68% Conservative: 29
Best Local Similarity: 64.17% Mismatches: 30
Query Match: 61.58% Indels: 8
DB: Gaps: 2

```

```

US-09-782-953-9 (1-197) x US-10-044-090-255 (1-2355)
QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValIleGlu 30
DB 181 AGCTCCCGATTGCTGCTGTGGCAACAGTGAATCTTACGCAAAAGTAAACCGAGGCC 240
QY 31 LysPheGlnGluValLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
DB 241 AAATTGAGTCCCTCTTTAGACGATGACAAAGACATACCTTTCACTATTAAAGAC 300
QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgIleGlu 70
DB 301 TTCAAAACGAGTCAAGATTAATCTTCAAGCAACCTTCTCCGACAGATCCAGGCTCCAG 360
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr 90
DB 361 CTGCATTAAGACTGAGCTTCTGGAAAGAAATGAATGATTATTTGCTCAACCTTACAC 420
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
DB 421 ATAGGAAGCTCA-----CACCTGCTCCGCCCAATCCAGACAAAGCATTTCTG 468
QY 111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130
DB 469 ATCTCCCTCCCGCTCTCTCCGCAAGTGGATGGAACAAGTGAATGCGACCCAGTC 528
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
DB 529 ATAACTATGATCTCTTATATGCCATCTCCAGCTGGGCCAGGGGAAAGATGATG 588
QY 151 HisAlaGlyThrGlnLysSerThrProSerValValHisValCysAspSerAspMetGlu 170
DB 589 CACGACGACACTGACACCACTCCACGCTGGTGTCTCATGTATGATGATGATCAAGAG 648
QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
DB 649 AAGGAGGAAGAAGGAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 708
QY 187 ArgArgProGlyLysProPro 193
DB 709 AGGAGGCCGAGATACACGCCG 729

```

RESULT 7  
 US-09-782-953-23  
 ; Sequence 23, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILTIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: US/09/782,953  
CURRENT APPLICATION NUMBER: 60/216,601  
PRIORITY FILING DATE: 2001-02-13  
PRIORITY FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 23  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(637)  
US-09-782-953-23

Alignment Scores:  
Pred. No.: 1,29e-61 Length: 720  
Score: 633.50 Matches: 127  
Percent Similarity: 77.01% Conservative: 17  
Best Local Similarity: 67.91% Mismatches: 32  
Query Match: 60.85% Indels: 11  
DB: 10 Gaps: 2

US-09-782-953-9 (1-197) x US-09-782-953-23 (1-720)

QY 11 SerThrLeuValAlaCysValAspValGluValPheThrAsnGlnGluValysGlu 30  
DB 83 ACTCAGCTTTTGGTTCGACGCTCCATGAAGCAGTGTTCAGGACGAGAGAGAGAA 142  
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
DB 143 AGATTGAAGCAGCTTTTCCACATCTATGATGACGAGTACTTTTTCAGCTGTTTAAAGC 202  
QY 51 PheArgValArgLysPheSerHisPheProLysSerAlaAlaArgAlaGlu 70  
DB 203 TTTAGAGAGTCAAGATAAATTTTCAGCAACCTGAAGCGGACGAGCGGAGAGAA 262  
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuLysLeuLysLeuLysLeu 90  
DB 263 CTCACGAAACAGACTTCAATGGCGAGAGCTAAGCTATTTTGCACAG----- 313  
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110  
DB 314 -----TCTATCTCTCGCGCCCGCAGCTGTCAAGCAGTTCCTC 352  
QY 111 IleSerProSerSerProValGlyTyrLysProLysSerAspAlaThrProVal 130  
DB 353 ATCTCCCTCCAGCCTCTCCCGCAGTGGGTGAAGCAGAGCGAGATGCGATGCTGTT 412  
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
DB 413 ATAAATATGATTTACTCTGTCTGTTTCCAAATGGGACGAGAGAGAAATATGAATT 472  
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170  
DB 473 CACCGGGAACACAGTCGACACCCAGCGTGTGTTCACTGTCTGTGAAGTGAACCTGA 532  
QY 171 GluGluGluAspProLysThrSerProLysProLysLysLeuGlnThrArgArgProGly 190  
DB 533 GAGGAG 589  
QY 191 LeuProSerValSerAsn 197  
DB 590 CTTCCGACCGCAGCTTGAAAT 610

RESULT 8

US-09-782-953-5  
Sequence 5, Application US/09782953  
Patent No. US20020150953A1

GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: US/09/782,953  
CURRENT APPLICATION NUMBER: 60/216,601  
PRIORITY FILING DATE: 2001-02-13  
PRIORITY FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 597  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(594)  
US-09-782-953-5

Alignment Scores:  
Pred. No.: 1,28e-61 Length: 597  
Score: 632.50 Matches: 121  
Percent Similarity: 77.13% Conservative: 24  
Best Local Similarity: 64.36% Mismatches: 34  
Query Match: 60.76% Indels: 9  
DB: 10 Gaps: 2

US-09-782-953-9 (1-197) x US-09-782-953-5 (1-597)

QY 11 SerThrLeuValAlaCysValAspValGluValPheThrAsnGlnGluValysGlu 30  
DB 31 AGCTCCCTGATTTGTTGTGCAACAGCATGATGTTTCAGGAAAGTGAAGACGGGCC 90  
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50  
DB 91 AAATTTGAATCTCTTTCAGAACATATGACAGGACAGCACCTTCCAGTATTTTAAAGAC 150  
QY 51 PheArgValArgLysPheSerHisPheProLysSerAlaAlaArgAlaGlu 70  
DB 151 TTCACAGCTGTCGAGATAAATTTTCAGCAACCTTATTCGACGCGATGCCAGGCTGCGG 210  
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuLysLeuLysLeuLysLeu 90  
DB 211 CTGCACAGCCGAGTCTCTGGGAGAGAAATGAAGTGTATTTTCTCAGACTTTACAC 270  
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110  
DB 271 ATAGGAAGTTCA-----CACCTGGCTCCGCCCAATCCCGACAAACAGTTCTC 318  
QY 111 IleSerProSerSerProValGlyTyrLysProLysSerAspAlaThrProVal 130  
DB 319 ATCTCCCTCCCGCCTCTCTCCCGTGGTGGAAACAGTAGAAGATGCCACCCCGTC 378  
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150  
DB 379 ATAAATATGATTTACTCTGTCTGTTTCCAAATGGGACGAGAGAGAGATGAAGTGA 438  
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170  
DB 439 CATGACGAGCAGAGACCCCACTCCAGTGTGTGTCACCGTGTGTGAGAGTGCACCAAG 498  
QY 171 GluGluGluAspProLys-----ThrSerProLysProLysLysLeuGln 185  
DB 499 AATGAGGAGAGAGAGAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558  
QY 186 ThrArgArgProGlyLeuProPro 193  
DB 559 ACACGGAGACCGGAGTACACACCG 582

RESULT 9

US-09-782-953-14





Db 499 AATGAGGAGGAGGAGGAGATGGAGAGATGAAGAGACCAAGCCCAAAATCATCCAG 558  
Qy 186 ThrArgArgProGlyLeuProPro 193  
Db 559 ACACGAGACCGGAGTACACACG 582  
RESULT 11  
US-09-880-107-3340  
; Sequence 3340, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3340  
; LENGTH: 2173  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833  
US-09-880-107-3340  
Alignment Scores:  
Pred. No.: 2,366-57 Length: 601.00 2173  
Score: 601.00 Matches: 115  
Percent Similarity: 79.77% Conservative: 23  
Best Local Similarity: 66.47% Mismatches: 27  
Query Match: 57.73% Indels: 8  
DB: 10 Gaps: 2  
US-09-782-953-9 (1-197) x US-09-880-107-3340 (1-2173)  
Qy 25 AsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAspGluCysValThr 44  
Db 43 AACAGATGGTGTATGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACC 102  
Qy 45 PheGlnLeuPheLysSerPheArgValArgLeuPheSerHisProLysSerAla 64  
Db 103 TTTGAGTATTTAAGACTTCAACAGAGTCAAGATTAACCTTACGACACCTTCTCCGCA 162  
Qy 65 AlaArgAlaArgLeuGluLeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyr 84  
Db 163 GCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTGGAAGGAATGAAGTTATAT 222  
Qy 85 PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProGln 104  
Db 223 TTTGCTCAGACCTTACATAGGAAGCTCA-----CACCTGGTCTCCGCCAAT 270  
Qy 105 ProAlaLysGlnPheLeuLysSerProSerSerProValGlyTyrLysProLeu 124  
Db 271 CCAGACNAGCAGTTTCGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAAGTG 330  
Qy 125 SerAspAlaThrProValLeuAsnTyrAspLeuLysTyrAlaValAlaLysLeuGlyPro 144  
Db 331 GAAGATGCCAGCCAGTCATAAATATGATCTCTTATATGCCATCTCCAGAGTGGGGCCA 390  
Qy 145 GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValValHisVal 164  
Db 391 GGGGAAAGATGATGAATTCACGACGAGGAGTACACCACTCCAGCGTGGTGGTCCATGTA 450  
Qy 165 CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys 180  
Db 165 CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys 180

Db 451 TGTGAGAGTGTATCAAGAGAGGAGGAGAGAGAGAAATGGAAGAATGAGGACCTAAG 510  
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProPro 193  
Db 511 CCAAAATTTATCCAGACCGAGGAGGCCGGAGTACACGCCG 549  
RESULT 12  
US-09-925-302-347  
; Sequence 347, Application US/09925302  
; Patent No. US20020044941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 347  
; LENGTH: 2358  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; OTHER INFORMATION: US-09-925-302-347  
Alignment Scores:  
Pred. No.: 1,796-55 Length: 584.50 2358  
Score: 584.50 Matches: 119  
Percent Similarity: 73.13% Conservative: 28  
Best Local Similarity: 59.20% Mismatches: 37  
Query Match: 56.15% Indels: 18  
DB: 10 Gaps: 3  
US-09-782-953-9 (1-197) x US-09-925-302-347 (1-2358)  
Qy 6 MetAspCysAspVal-----SerThrLeuValAlaCys 16  
Db 102 ATGACTGGAGATGGAGGAGGTGGACCTGCGAGGACCTGCCAGCCACCATCGCTGT 161  
Qy 17 ValValAspValGluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPhe 36  
Db 162 CACCTGGACCCCGCGGTGTCGTGGACGCGCTGTGCGGCGCCAAATTTGAGTCCCTCTT 221  
Qy 37 ArgThrTyrAspGluCysValThrPheGlnLeuPheLysSerPheArgValArgLeu 56  
Db 222 AGGACGTATGACAAGACACATCACCCTTTCAGTATTTTAAAGAGCTTCAACGAGTCAGAATA 281  
Qy 57 AsnPheSerHisProLysSerAlaAlaArgAlaArgLeuHisGluThrGlnPhe 76  
Db 282 AACTTCAGCAACCCCTTCTCCGACGAGATGCGAGCTCCAGCTGCATAGACTGAGTTT 341  
Qy 77 ArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAsp 96  
Db 342 CTGGGAAAGAAATGAAGTTATATTTTCTCAGACCTTACATAGGAAGCTCA----- 395  
Qy 97 LysLeuHisLeuAlaProGlnProAlaLysGlnPheLeuLysSerProSerSer 116  
Db 396 -----CACCTGGCTCCGCA-AATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCT 448  
Qy 117 ProProValGlyTyrLysProLysSerAspAlaThrProValLeuAsnTyrAspLeu 136  
Db 449 CCGCAGTGGGATGGAAACAGTGAAGATGCGACCCAGTCCATTAACATATGATCTCTTA 508  
Qy 137 TyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSer 156  
Db 509 TATGCCATCTCCAAGCTGGGGCCAGGGGAAAGATGATGAATTCACGACGAGTACAC 568  
Qy 157 ThrProSerValValHisValCysAspSerAspMetGluGluGluAspProLys 176  
Db 569 ACTCCAGCGTGTGGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628

Qy 177 Thrsr-----ProlyseProlylelleInThrArgProGlyLeuPro 192  
 Db 629 ATGAAGAATGAGAGACCTAAGCAAAATTTATCCAGACGAGCGCGAGTACACG 688  
 Qy 193 Pro 193  
 Db 689 CCG 691

## RESULT 13

US-09-782-953-1  
 : Sequence 1, Application US/09782953  
 : Patent No. US20020150953A1  
 : GENERAL INFORMATION:  
 : APPLICANT: WILLIAMS, R. SANDERS  
 : APPLICANT: ROTHERMEL, BEVERLY  
 : TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 : FILE REFERENCE: US/09/782,953  
 : CURRENT FILING DATE: 2001-02-13  
 : PRIOR FILING DATE: 2000-07-07  
 : PRIOR APPLICATION NUMBER: 60/216,601  
 : NUMBER OF SEQ ID NOS: 27  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 1  
 : LENGTH: 599  
 : TYPE: DNA  
 : ORGANISM: Mus musculus  
 : US-09-782-953-1

## Alignment Scores:

Pred. No.: 5,1e-54 Length: 599  
 Score: 564.00 Matches: 109  
 Percent Similarity: 81.10% Conservative: 24  
 Best Local Similarity: 66.46% Mismatches: 5  
 Query Match: 54.18% Indels: 1  
 DB: 10 Gaps: 1

US-09-782-953-9 (1-197) x US-09-782-953-1 (1-599)

Qy 11 SerThrLeuValAlaCysValValAlaPheValGluValPheThrArgInGluValValysGlu 30  
 Db 110 ACCTCCCTATTGCTGTGTGGCAACGATGCTTACGGAAGATGAGACCGAGCGC 169  
 Qy 31 LysPheGluGlyLeuPheArgThrTyraPheGluCysValThrPheGlnLeuPheLysSer 50  
 Db 170 AAATTGAAATCCCTCTTCGACACATGACACAGACACACCTTCCAGTATTTTAAGAC 229  
 Qy 51 PheArgArgValArgIleAsnPheserHisProlyseSerAlaAlaArgIleGlu 70  
 Db 230 TTCAAACGATCCGCGTAACTTACCAACCCCTTATTCGACGCGATCCAGCGTGGCG 289  
 Qy 71 LeuHisGluThrGlnPheArgGlyValysLeuLeuLeuTyraPheAlaGlnValGlnThr 90  
 Db 290 CTGCACAAACACGAGTCTCGGCGGAAGAAATGAAGTTGATTTCTCAGACTTTAC 349  
 Qy 91 ProGlnThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110  
 Db 350 ATAGAAAGTTCA-----CACCTGGCTCCGCCAAT-CCGACAAACAGTTCTC 396  
 Qy 111 IleserProPheSerSerProProValGlyTrpLysProIleserAspAlaThrProVal 130  
 Db 397 ATCTCCCTCCGCGCTCTCTCTCCGTTGGTGAACAGTAAGATGCGACCCCGTC 456  
 Qy 131 LeuAsnTyraPheLeuLeuTyraValAlaLysLeuGlyProGlyGluLysTyrgluLeu 150  
 Db 457 ATAAATTAAGATCTTTATATGCAATCTCCAACTCGGCGCAGAGACAGATGTGAAC 516  
 Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValLysAspSerAspMetGlu 170  
 Db 517 CATGACGACACACACCACTCCCAAGTGTGTGTCCACCTGTGTAGATGACCAAGAG 576

Qy 171 GluGluGluAsp 174  
 Db 577 AATGAGAGGAA 588

## RESULT 14

US-09-864-761-10388/C  
 : Sequence 10388, Application US/09864761  
 : Patent No. US20020048763A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Penn, Sharon G.  
 : APPLICANT: Rank, David R.  
 : APPLICANT: Hanzel, David K.  
 : APPLICANT: Chen, Wensheng  
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 : FILE REFERENCE: Aecm1ca-X-1  
 : CURRENT FILING DATE: US/09/864,761  
 : PRIOR FILING DATE: 2001-05-23  
 : PRIOR APPLICATION NUMBER: US 60/180,312  
 : PRIOR FILING DATE: 2000-02-04  
 : PRIOR APPLICATION NUMBER: US 60/207,456  
 : PRIOR FILING DATE: 2000-05-26  
 : PRIOR APPLICATION NUMBER: US 09/632,366  
 : PRIOR FILING DATE: 2000-08-03  
 : PRIOR APPLICATION NUMBER: GB 24263,6  
 : PRIOR FILING DATE: 2000-10-04  
 : PRIOR APPLICATION NUMBER: US 60/236,359  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: PCT/US01/00666  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00667  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00664  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00669  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00665  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00668  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00663  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00662  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00661  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: PCT/US01/00670  
 : PRIOR FILING DATE: 2001-01-30  
 : PRIOR APPLICATION NUMBER: US 60/234,687  
 : PRIOR FILING DATE: 2000-09-21  
 : PRIOR APPLICATION NUMBER: US 09/608,408  
 : PRIOR FILING DATE: 2000-06-30  
 : PRIOR APPLICATION NUMBER: US 09/774,203  
 : PRIOR FILING DATE: 2001-01-29  
 : NUMBER OF SEQ ID NOS: 49117  
 : SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 : SEQ ID NO 10388  
 : LENGTH: 412  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens

## FEATURE:

OTHER INFORMATION: MAP TO AF000054.1  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67  
 OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67  
 US-09-864-761-10388

Alignment Scores:  
 Pred. No.: 2e-16

Length: 412

Score: 223.00 Matches: 42  
Percent Similarity: 89.66% Conservative: 10  
Best Local Similarity: 72.41% Mismatches: 6  
Query Match: 21.42% Indels: 0  
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x US-09-864-761-10388 (1-412)

Qy 31 LysPheGluGlyLeuPheArgThrTyraPspGluCysValThrPheGlnLeuPheLysSer 50  
Db 407 AAATTGGTCCCTTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 348  
Qy 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
Db 347 TTCAACGAGTCAGATAAACTTTCAGAACCCCTTCCGACGACATGCCAGGCTCCAG 288

Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyPheAlaGlnVal 88  
Db 287 CTGCATAAGACTGAGTTCTGGGAAAGGAATGAAGTTATATTTTCTCAGGTG 234

RESULT 15

US-09-864-761-2064/c  
Sequence 2064, Application US/09864761  
Patent No. US2002048763A1  
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aomicra-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 2064

LENGTH: 446  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AF000122.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4  
US-09-864-761-2064

Alignment Scores:  
Pred. No.: 2.23e-16 Length: 446  
Score: 223.00 Matches: 42  
Percent Similarity: 89.66% Conservative: 10  
Best Local Similarity: 72.41% Mismatches: 6  
Query Match: 21.42% Indels: 0  
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x US-09-864-761-2064 (1-446)

Qy 31 LysPheGluGlyLeuPheArgThrTyraPspGluCysValThrPheGlnLeuPheLysSer 50  
Db 441 AAATTGGTCCCTTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 382  
Qy 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70  
Db 381 TTCAACGAGTCAGATAAACTTTCAGAACCCCTTCCGACGACATGCCAGGCTCCAG 322  
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyPheAlaGlnVal 88  
Db 321 CTGCATAAGACTGAGTTCTGGGAAAGGAATGAAGTTATATTTTCTCAGGTG 268

Search completed: December 15, 2002, 00:06:30  
Job time : 37.8887 secs

**THIS PAGE BLANK (USPTO)**

```

RESULT 2
US-09-782-953-10
; Sequence 10, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-10

Query Match      100.0%; Score 1041; DB 10; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.2e-94;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAPSMDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
DB 1 MPAPSMDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 61 PKSAARARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
DB 61 PKSAARARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 121 WKPISDATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
DB 121 WKPISDATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 181 PKIITRRPGLPPSVSN 197
DB 181 PKIITRRPGLPPSVSN 197

RESULT 3
US-09-782-953-18
; Sequence 18, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-18

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 95.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 65
DB 1 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 4
US-09-782-953-19
; Sequence 19, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 65
DB 1 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 5
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 95.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 65
DB 1 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

```

```

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 4
US-09-782-953-19
; Sequence 19, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 65
DB 1 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 5
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 65
DB 1 MDCDVSTLVACVVDVEVFTNQEVKEFGLFRTYDECVTQQLFKSFRRVRINFESH 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVWGKPI 120
QY 126 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 185
DB 121 DATPVLYDLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKP 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:49 ; Search time 6.31197 Seconds  
(without alignments)  
506.931 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041  
Sequence: 1 MPAPSMDCVSTLVACVVDV.....SPKPIQTRRPGLPSPSVN 197

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_Aa:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*

9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*

12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	US-09-782-953-9	Sequence 9, Appl
2	1041	100.0	197	US-09-782-953-10	Sequence 10, Appl
3	980	94.1	192	US-09-782-953-18	Sequence 18, Appl
4	980	94.1	192	US-09-782-953-19	Sequence 19, Appl
5	658.5	63.3	241	US-09-782-953-21	Sequence 21, Appl
6	658.5	63.3	241	US-09-782-953-22	Sequence 22, Appl
7	641	61.6	197	US-09-782-953-12	Sequence 12, Appl
8	641	61.6	197	US-09-782-953-13	Sequence 13, Appl
9	633.5	60.9	212	US-09-782-953-24	Sequence 24, Appl
10	633.5	60.9	212	US-09-782-953-25	Sequence 25, Appl
11	632.5	60.8	198	US-09-782-953-6	Sequence 6, Appl
12	632.5	60.8	198	US-09-782-953-7	Sequence 7, Appl
13	622	59.8	197	US-09-782-953-15	Sequence 15, Appl
14	622	59.8	197	US-09-782-953-16	Sequence 16, Appl
15	613.5	58.9	198	US-09-782-953-4	Sequence 3, Appl
16	613.5	58.9	198	US-09-782-953-3	Sequence 4, Appl
17	254.5	24.4	142	US-09-925-302-790	Sequence 790, App
18	219	21.0	58	US-09-864-761-35379	Sequence 35379, A
19	219	21.0	58	US-09-864-761-43076	Sequence 43076, A

20	174	16.7	56	10	US-09-864-761-34111	Sequence 34111, A
21	90	8.6	526	10	US-09-801-368-362	Sequence 362, App
22	85	8.2	4019	10	US-09-738-973-425	Sequence 425, App
23	82	7.9	2783	10	US-09-816-669A-14	Sequence 14, Appl
24	81	7.8	559	10	US-09-854-549-7	Sequence 7, Appl
25	81	7.8	559	10	US-09-856-561-6	Sequence 6, Appl
26	79	7.6	985	9	US-09-978-285A-211	Sequence 211, App
27	79	7.6	985	9	US-09-978-697-211	Sequence 211, App
28	79	7.6	985	9	US-09-978-697-211	Sequence 211, App
29	78.5	7.5	932	10	US-09-754-997A-4	Sequence 4, Appl
30	78.5	7.5	1252	10	US-09-754-997A-2	Sequence 2, Appl
31	77	7.4	1395	10	US-09-858-664A-11	Sequence 11, Appl
32	76.5	7.3	330	10	US-09-794-591-2	Sequence 2, Appl
33	75	7.2	976	10	US-09-969-528-2	Sequence 2, Appl
34	75	7.2	2441	12	US-09-109-886-8	Sequence 8, Appl
35	74	7.1	433	10	US-09-925-300-1467	Sequence 1467, App
36	73	7.0	433	10	US-09-764-864-977	Sequence 977, App
37	72.5	7.0	228	10	US-09-864-761-37276	Sequence 37276, A
38	72.5	7.0	282	10	US-09-864-761-35386	Sequence 35386, A
39	72.5	7.0	282	10	US-09-864-761-37704	Sequence 37704, A
40	72.5	7.0	598	10	US-09-853-386-91	Sequence 91, Appl
41	72.5	7.0	598	10	US-09-853-386-98	Sequence 98, Appl
42	72.5	7.0	748	10	US-09-864-761-43244	Sequence 43244, A
43	72	6.9	662	10	US-09-799-777-53	Sequence 53, Appl
44	72	6.9	662	10	US-09-771-161A-164	Sequence 164, App
45	72	6.9	743	10	US-09-771-161A-254	Sequence 254, App

## ALIGNMENTS

RESULT 1

US-09-782-953-9

Sequence 9, Application US/09782953

Parent No. US20020150953A1

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

FILE REFERENCE: US20020150953A1

CURRENT APPLICATION NUMBER: US/09/782,953

PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 197

TYPE: PRT

ORGANISM: Mus musculus

US-09-782-953-9

Query Match 100.0%, Score 1041, DB 10, Length 197,  
Best Local Similarity 100.0%, Pred. No. 1.2e-94,  
Matches 197, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY	1	MPAPSMDCVSTLVACVVDVFTNOEVEKEFGLFRITDECVTFOLPFKSFRVRINFSH	60
DB	1	MPAPSMDCVSTLVACVVDVFTNOEVEKEFGLFRITDECVTFOLPFKSFRVRINFSH	60
QY	61	PKSAARIEIHEHETORFGKKLKLYFAOVOTPETDGLHAPPOPAKOFILSPSSPPVG	120
DB	61	PKSAARIEIHEHETORFGKKLKLYFAOVOTPETDGLHAPPOPAKOFILSPSSPPVG	120
QY	121	WKISATPTLVANPLYAVALKGLPGEKYLHAATESPSVYVHVCSDMEEDPKTSPK	180
DB	121	WKISATPTLVANPLYAVALKGLPGEKYLHAATESPSVYVHVCSDMEEDPKTSPK	180
QY	181	PKIOTRRPGLPSPSVN 197	
DB	181	PKIOTRRPGLPSPSVN 197	

; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: ParentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-21

Query Match 63.3%; Score 658.5; DB 21; Length 241;  
Best Local Similarity 70.1%; Pred. No. 3.1e-58;  
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;  
Qy 11 STLACVVDVEFTNQEVKEKEGEGFRYDECVTQLEKSPRRVRINFSPKSAARARIE 70  
Db 47 TSLFACSVHEAVFEAREOKERPEALFTIYDDQVTFQLEKSPRRVRINFSPKSAARARIE 106  
Qy 71 LHETQFRGKGLKLYPAQVQTPETDGDGLHLAPPQPAKQFLISPPSSPPVGVKPISDATPV 130  
Db 107 LHETDFNGKGLKLYPAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVGVKQSEDAMPV 166  
Qy 131 LNYDLLYAVAKLPGCKYELHAGTSTPSVVHVCDSDMEDEEDPKTSPKPIQTRRPG 190  
Db 167 INYDLLCAVSKLPGCKYELHAGTSTPSVVHVHVESETEEEETK-NPKQKIAQTRRPD 225  
Qy 191 LPPSVSN 197  
Db 226 PPTAALN 232

Search completed: December 11, 2002, 11:49:34  
Job time : 135.874 secs

Db 249 LPPSVSN 255

RESULT 12

US-09-575-580B-5  
; Sequence 5, Application US/09575580B  
; GENERAL INFORMATION:  
; APPLICANT: McKean, F.  
; APPLICANT: McKean, K.  
; APPLICANT: Ryeon, S.  
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN  
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO  
; FILE REFERENCE: HMV-048.01  
; CURRENT APPLICATION NUMBER: US/09/575,580B  
; CURRENT FILING DATE: 2000-05-22  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-575-580B-5

Query Match

Best Local Similarity 84.7%; Score 682; DB 19; Length 242;  
Best Local Similarity 91.8%; Pred. No. 4,6e-81;  
Matches 166; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 70

Db 57 NSLFACVHOSVFEESKEKEFGLFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 116

QY 71 LHETORFGKLLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 117 LHETORFGKLLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 176

QY 131 LNYDLVAAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDDPKTSPPKXIQTTRPG 190

Db 177 LNYDLVAAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDDPKTSPPKXIQTTRPG 236

QY 191 LPP 193

Db 237 LPP 239

RESULT 13

PCT-US02-17382-133  
; Sequence 133, Application PC/TUS0217382  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-062  
; CURRENT APPLICATION NUMBER: PCT/US02/17382  
; CURRENT FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/326,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 234  
; SOFTWARE: PatentIn Version 3.1  
; SEQ ID NO 133  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-17382-133

Query Match

Best Local Similarity 63.3%; Score 658.5; DB 1; Length 241;  
Best Local Similarity 70.1%; Pred. No. 3.1e-58;  
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 70

Db 47 TSLFACVHAEVFEAREQKRFELFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 106

QY 71 LHETORFGKLLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 107 LHETDFNGQLLKYFAOVQMSGEVRDYSYLLPQPVKQFLISPPASPPVGMKQSEDMAPV 166

QY 131 LNYDLVAAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDDPKTSPPKXIQTTRPG 190

Db 167 INVDLCANVSKLGPGEKYEELHAGTSTPSVVVHVCSETEEBEETK-NPKOKIAQTRRPD 225

QY 191 LPPSVSN 197

Db 226 PPTAALN 232

RESULT 14

US-09-614-474-11  
; Sequence 11, Application US/09614474  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; APPLICANT: Streeter, David G.  
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
; FILE REFERENCE: PC-0013 US  
; CURRENT APPLICATION NUMBER: US/09/614,474  
; CURRENT FILING DATE: 2000-07-11  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PERL Program  
; SEQ ID NO 11  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc:feature  
; OTHER INFORMATION: Incyte ID No: 96017919  
US-09-614-474-11

Query Match

Best Local Similarity 63.3%; Score 658.5; DB 20; Length 241;  
Best Local Similarity 70.1%; Pred. No. 3.1e-58;  
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 70

Db 47 TSLFACVHAEVFEAREQKRFELFTYDVCYTFQLFKSFRRVRINFSHPKSAARARIE 106

QY 71 LHETORFGKLLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 107 LHETDFNGQLLKYFAOVQMSGEVRDYSYLLPQPVKQFLISPPASPPVGMKQSEDMAPV 166

QY 131 LNYDLVAAVAKLGPGEKYEELHAGTSTPSVVVHVCSDMEEDDPKTSPPKXIQTTRPG 190

Db 167 INVDLCANVSKLGPGEKYEELHAGTSTPSVVVHVCSETEEBEETK-NPKOKIAQTRRPD 225

QY 191 LPPSVSN 197

Db 226 PPTAALN 232

RESULT 15

US-09-782-953-21  
; Sequence 21, Application US/09782953  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07



```

Query Match      86.9%; Score 905; DB 24; Length 234;
Best Local Similarity 91.4%; Pred. No. 2e-83;
Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 11 STLVAQVVDVEFTNOEVEKKEFGLFRTYDCCVTFOLPKSPRRVRINFSHPKSAARAE 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 NSLFACNVHOSVFGEESKEFGLFRTYDCCVTFOLPKSPRRVRINFSHPKSAARAE 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHETQFRGKKLKYFAQVQTPETDGDKLHLAPPQAKOFLISPPSSPPVGWKPISDATPV 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 LHETQFRGKKLKYFAQVQTPETDGDKLHLAPPQAKOFLISPPSSPPVGWQPIN DATPV 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 131 LNYDLLYAVAKLGPGEKVELHAGHAGTSTPSVVHVCDSDMEEDPKTSPKPKIIQTRRPG 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 LNYDLLYAVAKLGPGEKVELHAGHAGTSTPSVVHVCDSDIEEDPKTSPKPKIIQTRRPG 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 191 LPPSVSN 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 LPPSVSN 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-60-142-678-1
; Sequence 1, Application US/60142678
; GENERAL INFORMATION:
; APPLICANT: Tang, Tom, Y.
; TITLE OF INVENTION: HUMAN PROTEINS INVOLVED IN DETOXIFICATION
; FILE REFERENCE: PF-0711 P
; CURRENT APPLICATION NUMBER: US/60/142,678
; CURRENT FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc features
; OTHER INFORMATION: Incyte Clone No.: 3751586
US-60-142-678-1

Query Match      86.9%; Score 905; DB 27; Length 234;
Best Local Similarity 91.4%; Pred. No. 2e-83;
Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 11 STLVAQVVDVEFTNOEVEKKEFGLFRTYDCCVTFOLPKSPRRVRINFSHPKSAARAE 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 NSLFACNVHOSVFGEESKEFGLFRTYDCCVTFOLPKSPRRVRINFSHPKSAARAE 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHETQFRGKKLKYFAQVQTPETDGDKLHLAPPQAKOFLISPPSSPPVGWKPISDATPV 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 LHETQFRGKKLKYFAQVQTPETDGDKLHLAPPQAKOFLISPPSSPPVGWQPIN DATPV 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 131 LNYDLLYAVAKLGPGEKVELHAGHAGTSTPSVVHVCDSDMEEDPKTSPKPKIIQTRRPG 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 LNYDLLYAVAKLGPGEKVELHAGHAGTSTPSVVHVCDSDIEEDPKTSPKPKIIQTRRPG 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 191 LPPSVSN 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 LPPSVSN 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-10-104-047-2216
; Sequence 2216, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:

```

APPLICANT: Edwards, Carla M.  
APPLICANT: Streeter, David G.  
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
FILE REFERENCE: PC-0013 US  
CURRENT APPLICATION NUMBER: US/09/614,474  
CURRENT FILING DATE: 2000-07-11  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PERL Program  
SEQ ID NO 10  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: g1435040  
US-09-614-474-10

Query Match 94.1%; Score 980; DB 20; Length 192;  
Best Local Similarity 96.4%; Pred. No. 3.2e-91;  
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 65  
DB 1 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 60  
QY 66 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 125  
DB 61 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 120  
QY 126 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEBEDPKTSKPKIIQ 185  
DB 121 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEBEDPKTSKPKIIQ 180  
QY 186 TRRPGLPSPSVN 197  
DB 181 TRRPGLPSPSVN 192

RESULT 6  
US-09-782-953-18  
Sequence 18, Application US/09782953  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: UTSID:674PZ1  
CURRENT APPLICATION NUMBER: US/09/782,953  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/216,601  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-782-953-18

Query Match 94.1%; Score 980; DB 21; Length 192;  
Best Local Similarity 96.4%; Pred. No. 3.2e-91;  
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 65  
DB 1 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 60  
QY 66 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 125  
DB 61 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 120  
QY 126 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEBEDPKTSKPKIIQ 185  
DB 121 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEBEDPKTSKPKIIQ 180

DB 121 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEBEDPKTSKPKIIQ 180  
QY 186 TRRPGLPSPSVN 197  
DB 181 TRRPGLPSPSVN 192

RESULT 7  
US-09-782-953-19  
Sequence 19, Application US/09782953  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, R. SANDERS  
APPLICANT: ROTHERMEL, BEVERLY  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
FILE REFERENCE: UTSID:674PZ1  
CURRENT APPLICATION NUMBER: US/09/782,953  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/216,601  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 19  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-782-953-19

Query Match 94.1%; Score 980; DB 21; Length 192;  
Best Local Similarity 96.4%; Pred. No. 3.2e-91;  
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 65  
DB 1 MDCDVSTLVACVVDVEFTNOEYKEKFGGLFRITYDCVTFOLFKSPRRVRINFSHPKSA 60  
QY 66 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 125  
DB 61 RARIELHETOPRGKKLKYFAQVOTPETDGDKLHAPPOPAKQFLISPPSSPPVGMKPI 120  
QY 126 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEBEDPKTSKPKIIQ 185  
DB 121 DATPVNLYDLIYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEBEDPKTSKPKIIQ 180  
QY 186 TRRPGLPSPSVN 197  
DB 181 TRRPGLPSPSVN 192

RESULT 8  
US-10-030-613-1  
Sequence 1, Application US/10030613  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: TANG, Y. TOM  
APPLICANT: YUE, HENRY  
TITLE OF INVENTION: HUMAN PROTEINS INVOLVED IN DETOXIFICATION  
FILE REFERENCE: PF-0711 PCT  
CURRENT APPLICATION NUMBER: US/10/030,613  
CURRENT FILING DATE: 2002-01-03  
PRIOR APPLICATION NUMBER: 60/142,678  
PRIOR FILING DATE: 1999-07-07  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PERL Program  
SEQ ID NO 1  
LENGTH: 234  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: 3751586CD1  
US-10-030-613-1

Qy	61	PKSAARARIELHETQPRGKKLKYFAOVQTPETDGDKLHLAPQPAKQFLISPPSSPPVG	120
Db	61	PKSAARARIELHETQPRGKKLKYFAOVQTPETDGDKLHLAPQPAKQFLISPPSSPPVG	120
Qy	121	WKPISDATPVNLVDLLYAVAKLGPGEKYLHAGTESTPSVVHVCDSDMEEEDPKTSPK	180
Db	121	WKPISDATPVNLVDLLYAVAKLGPGEKYLHAGTESTPSVVHVCDSDMEEEDPKTSPK	180
Qy	181	PKIIQTRRPGLPSPSVSN	197
Db	181	PKIIQTRRPGLPSPSVSN	197

```

RESULT 2
US-09-782-953-10
; Sequence 10, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSID:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 10
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-10

```

Query Match	100.0%;	Score 1041;	DB 21;	Length 197;
Best Local Similarity	100.0%;	Pred. No. 2e-97;		
Matches 197;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPAPSMDCDVSTLVACVVDVEVFNQEVKEFGLRTYDECVTQFQFKSFRVRINFSH	60	
Db	1	MPAPSMDCDVSTLVACVVDVEVFNQEVKEFGLRTYDECVTQFQFKSFRVRINFSH	60	
Qy	61	PKSAARARIELHETQFRGKKLKYFAQVQTPETDGDKHLAPPQAKQLIGPPSSPPVG	120	
Db	61	PKSAARARIELHETQFRGKKLKYFAQVQTPETDGDKHLAPPQAKQLIGPPSSPPVG	120	
Qy	121	WKPISDATPVNLDDLKYAVAKLGGKGYELHAGHTESTPSVVHVCDSDMEEEDPKTSPK	180	
Db	121	WKPISDATPVNLDDLKYAVAKLGGKGYELHAGHTESTPSVVHVCDSDMEEEDPKTSPK	180	
Qy	181	PKIIOTRRPGJPPPSVSN	197	
Db	181	PKIIOTRRPGJPPPSVSN	197	

```

RESULT 3
PCT-US02-17382-132
    ; Sequence 132, Application PC/TUS0217382
    ; GENERAL INFORMATION:
    ; APPLICANT: EXELIXIS, INC.
    ; TITLE OF INVENTION: MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
    ; FILE REFERENCE: EX02-062
    ; CURRENT APPLICATION NUMBER: PCT/US02/17382
    ; CURRENT FILING DATE: 2002-06-05
    ; PRIOR APPLICATION NUMBER: US 60/236,076
    ; PRIOR FILING DATE: 2001-06-05
    ; PRIOR APPLICATION NUMBER: US 60/328,605
    ; PRIOR FILING DATE: 2001-10-10
    ; PRIOR APPLICATION NUMBER: US 60/357,253
    ; PRIOR FILING DATE: 2002-02-15
    ; NUMBER OF SEQ ID NOS: 234
    ; SOFTWARE: PatentIn version 3.1
    ; SEQ ID NO 132

```

```

; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
:PCT-US02-17382-132

Query March          98.0%; Score 1020; DB 1; Length 197;
Best Local Similarity 97.5%; Pred. No. 2.7e-95;
Matches 197; Conservative 5; Mismatches 0; Indels 0; Caps 0;

```

RESULT 4  
 US-09-575-580B-6  
 ; Sequence 6, Application US/09575580B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McKeon, F.  
 ; APPLICANT: Kayako, K.  
 ; APPLICANT: Ryecroft, S.  
 ; TITLE OF INVENTION: CALCIOPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN  
 ; FILE OF INVENTION: USES AND REAGENTS RELATED THERETO  
 ; FILE REFERENCE: HMV-048.01  
 ; CURRENT APPLICATION NUMBER: US/09/575,580B  
 ; CURRENT FILING DATE: 2000-05-22  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-575-580B-6

Query Match	95.4%	Score	993;	DB	19;	Length	192;
Best Local Similarity	97.4%;	Pred. No.	1.5e-92;				
Matches	187;	Conservative	5;	Mismatches	0;	Indels	0;
						Gaps	0;
QY	6	MDCDVSTLVACVVDVEVFTNQEVKEKEGEGFRITVDCVQTQOLFKFSFRVRVIRNFNSHPKSA	65				
Db	1	MDCDVSTLVACVVDVEVFTNQEVKEKEGEGFRITVDDCVQTQOLFKFSFRVRVIRNFNSPKSA	60				
QY	66	RARIELHETQFRGKKLKLFAQVQTPTDGDGKLHLAPPQAKQFLISPPSPPPVGMKPIIS	125				
Db	61	RARIELHETQFRGKKLKLFAQVQTPTDGDGKLHLAPPQAKQFLISPPSPPPVGMQPIN	120				
QY	126	DATPVNLNLLLYAVAKLGPGKEYELHAGTSTSPVVHVCDSDMEEREEDPKTSPKPKIIQ	185				
Db	121	DATPVNLNLLLYAVAKLGPGKEYELHAGTSTSPVVHVCDSDIEEEDPKTSPKPKIIQ	180				
QY	186	TRRPGLPSPSVN	197				
Db	181	TRRPGLPSPSVN	192				

RESULT 5  
US-09-614-474-10  
; Sequence 10, Application US/09614474  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 135.874 Seconds  
(without alignments)

934.784 Million cell updates/sec

Title: US-09-782-953-9

Sequence: 1 MPAPSMDCDVSTLVACVVDV.....SPKPKIQTRRPGLPPSVSN 197

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 reqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/prodata/1/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/paa/US07\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
7: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
8: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
9: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
10: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
11: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
12: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
13: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
14: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
15: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
16: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
17: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
18: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
19: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
20: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
21: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
22: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
23: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
24: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
25: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
26: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*  
27: /cgn2\_6/prodata/1/paa/US09\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	21	US-09-782-953-9
2	1041	100.0	197	21	US-09-782-953-10
3	1020	98.0	197	1	PCT-US07-1782-132
4	993	95.4	192	19	US-09-575-5808-6
5	980	94.1	192	20	US-09-514-474-10
6	980	94.1	192	21	US-09-782-953-18

7	980	94.1	192	21	US-09-782-953-19	Sequence 19, Appl
8	905	86.9	234	24	US-10-030-613-1	Sequence 1, Appl1
9	905	86.9	234	27	US-60-142-678-1	Sequence 1, Appl1
10	905	86.9	243	25	US-10-104-047-2216	Sequence 216, Ap
11	905	86.9	255	20	US-09-614-474-2	Sequence 2, Appl1
12	882	84.7	242	19	US-09-575-5808-5	Sequence 5, Appl1
13	658.5	63.3	241	1	PCT-US02-1782-133	Sequence 133, App
14	658.5	63.3	241	20	US-09-614-474-11	Sequence 11, Appl
15	658.5	63.3	241	21	US-09-782-953-11	Sequence 22, Appl
16	658.5	63.3	241	21	US-09-782-953-12	Sequence 24, Appl
17	646	62.1	236	19	US-09-575-5808-24	Sequence 8, Appl1
18	641	61.6	197	19	US-09-575-5808-8	Sequence 12, Appl
19	641	61.6	197	21	US-09-782-953-12	Sequence 13, Appl
20	641	61.6	197	21	US-09-782-953-13	Sequence 14, App
21	641	61.6	197	26	US-10-247-671-174	Sequence 174, Appl
22	641	61.6	197	27	US-09-782-953-14	Sequence 24, Appl
23	633.5	60.9	212	21	US-09-782-953-25	Sequence 25, Appl
24	633.5	60.9	212	21	US-09-782-953-6	Sequence 6, Appl1
25	632.5	60.8	198	21	US-09-782-953-7	Sequence 7, Appl1
26	632.5	60.8	198	21	US-10-104-047-2892	Sequence 2892, Ap
27	626.5	60.2	252	25	US-09-782-953-15	Sequence 15, Appl
28	622	59.8	197	21	US-09-782-953-16	Sequence 16, Appl
29	622	59.8	197	21	US-09-782-953-16	Sequence 4, Appl1
30	613.5	58.9	198	19	US-09-575-5808-4	Sequence 3, Appl1
31	613.5	58.9	198	21	US-09-782-953-3	Sequence 4, Appl1
32	613.5	58.9	198	21	US-09-782-953-4	Sequence 7, Appl1
33	596	57.3	170	19	US-09-575-5808-7	Sequence 6466, Ap
34	476.5	45.8	142	21	US-09-782-953-15	Sequence 6466, Ap
35	476.5	45.8	142	26	US-10-235-926-6466	Sequence 6466, Ap
36	404	38.8	292	20	US-09-614-150-41193	Sequence 14193, A
37	404	38.8	292	20	US-09-619-049-1440	Sequence 1440, Ap
38	404	38.8	292	27	US-60-167-334-118	Sequence 718, App
39	404	38.8	292	27	US-60-171-627-2109	Sequence 2109, App
40	404	38.8	292	27	US-60-173-386-180	Sequence 684, App
41	404	38.8	292	27	US-60-175-871-764	Sequence 764, App
42	404	38.8	292	27	US-60-184-775-692	Sequence 692, App
43	404	38.8	292	27	US-60-191-637-40818	Sequence 40818, A
44	404	38.8	292	27	US-60-191-700-744	Sequence 744, App
45	400.5	38.5	164	27	US-60-140-956-1718	Sequence 1718, Ap

#### ALIGNMENTS

RESULT 1

US-09-782-953-9 Application US/09782953

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

TITLE OF INVENTION: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

FILE REFERENCE: US02:674P21

CURRENT APPLICATION NUMBER: US/09/782,953

PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 197

TYPE: PRT

ORGANISM: Mus musculus

US-09-782-953-9

Query Match 100.0%; Score 1041; DB 21; Length 197;

Best Local Similarity 100.0%; Pred No. 26-97; 0; Indels 0; Gaps 0;

Matches 197; Conservative 0; Mismatches 0;

Qy 1 MPAPSMDCDVSTLVACVVDVFTNOEVEKKEKFKGLPTTYDECYTFOLFKSFRFRVIRNF5H 60

Db 1 MPAPSMDCDVSTLVACVVDVFTNOEVEKKEKFKGLPTTYDECYTFOLFKSFRFRVIRNF5H 60

**THIS PAGE BLANK (USPTO)**

```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61352
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61352

Query Match          59.8%; Score 622; DB 5; Length 197;
Best Local Similarity 63.6%; Pred. No. 1.4e-49;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLVACVVDVEFTNQVEKEKEGELFRTYDECVTFLFKSFRVRINFSHPKSAARARIE 70
DB 11 SATTACHIDPVPVFDGLCRARFESLFRITYDKDITFOYFKSFKVRINFSNPFSAADARLQ 70
QY 71 LHETQFRGKKLKYFAOVQTPETDGDKLHAPQPAKQFLISPPSPPVGKPKISDATPV 130
DB 71 LHKTEFLGKEMKLYFAQTLHGSS---HLAPPNPKQFLISPPASPPVGKQVEDATPV 126
QY 131 LVNYDLVAVAKLGPEKYEELHAGTESPTSVVHVCDSDMEBEDPKTS---PKPKIIT 186
DB 127 INVLLVLAISKLGPGEKYEELHAATDTTPSVVHVCDSDQKEEEDERMRMRPKKIIT 186
QY 187 QTRRPGLLP 193
DB 187 RRPPEYTP 193

RESULT 14
US-09-724-676-61356
; Sequence 61356, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61356
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61356

Query Match          57.4%; Score 598; DB 5; Length 200;
Best Local Similarity 61.9%; Pred. No. 2.2e-47;
Matches 117; Conservative 23; Mismatches 37; Indels 12; Gaps 3;

QY 9 DVSTLVACVVDVEFTNQVEKEKEGELFRTYDECVTFLFKSFRVRINFSHPKSAARAR 68
DB 16 DVAPFPICTIG---RKANWMAKFESLFRITYDKDITFOYFKSFKVRINFSNPFSAADAR 71
QY 69 IELEHTQFRGKKLKYFAOVQTPETDGDKLHAPQPAKQFLISPPSPPVGKPKISDAT 128
DB 72 LQHKTEFLGKEMKLYFAQTLHGSS---HLAPPNPKQFLISPPASPPVGKQVEDAT 127
QY 129 PVINYDLVAVAKLGPEKYEELHAGTESPTSVVHVCDSDMEBEDPKTS---PKPKI 184
DB 128 PVINYDLVLAISKLGPGEKYEELHAATDTTPSVVHVCDSDQKEEEDERMRMRPKKI 187
QY 185 QTRRPGLLP 193
DB 188 QTRRPEYTP 196

```

```

RESULT 15
US-09-724-676A-61356
; Sequence 61356, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61356
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61356

Query Match          57.4%; Score 598; DB 5; Length 200;
Best Local Similarity 61.9%; Pred. No. 2.2e-47;
Matches 117; Conservative 23; Mismatches 37; Indels 12; Gaps 3;

QY 9 DVSTLVACVVDVEFTNQVEKEKEGELFRTYDECVTFLFKSFRVRINFSHPKSAARAR 68
DB 16 DVAPFPICTIG---RKANWMAKFESLFRITYDKDITFOYFKSFKVRINFSNPFSAADAR 71
QY 69 IELEHTQFRGKKLKYFAOVQTPETDGDKLHAPQPAKQFLISPPSPPVGKPKISDAT 128
DB 72 LQHKTEFLGKEMKLYFAQTLHGSS---HLAPPNPKQFLISPPASPPVGKQVEDAT 127
QY 129 PVINYDLVAVAKLGPEKYEELHAGTESPTSVVHVCDSDMEBEDPKTS---PKPKI 184
DB 128 PVINYDLVLAISKLGPGEKYEELHAATDTTPSVVHVCDSDQKEEEDERMRMRPKKI 187
QY 185 QTRRPGLLP 193
DB 188 QTRRPEYTP 196

```

Search completed: December 11, 2002, 11:50:15  
 Job time : 11.6273 secs

; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61358  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61358

Query Match 61.6%; Score 641; DB 5; Length 197;  
Best Local Similarity 64.2%; Pred. No. 2.6e-51;  
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDCVTFQLEKSPRRVRINFSPKSAARARIE 70  
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPFSAADARLQ 70  
Qy 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAPAKOFLISPPSPVGVGKPISDATPV 130  
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVGKQVEDATPV 126  
Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186  
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186  
Qy 187 RRPGLPP 193  
Db 187 RRPEYTP 193

## RESULT 10

US-09-724-676A-61359  
; Sequence 61359, Application US/09724676A  
; GENERAL INFORMATION:

; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724.676A  
; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 61359  
; LENGTH: 197

; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-724-676A-61359

Query Match 61.6%; Score 641; DB 5; Length 197;  
Best Local Similarity 64.2%; Pred. No. 2.6e-51;  
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDCVTFQLEKSPRRVRINFSPKSAARARIE 70  
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPFSAADARLQ 70  
Qy 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAPAKOFLISPPSPVGVGKPISDATPV 130  
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVGKQVEDATPV 126  
Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186  
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186  
Qy 187 RRPGLPP 193  
Db 187 RRPEYTP 193

## RESULT 11

US-09-724-676A-61367

; Sequence 61367, Application US/09724676A  
; GENERAL INFORMATION:

; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724.676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61367  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61367

Query Match 61.6%; Score 641; DB 5; Length 197;  
Best Local Similarity 64.2%; Pred. No. 2.6e-51;  
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDCVTFQLEKSPRRVRINFSPKSAARARIE 70  
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPFSAADARLQ 70  
Qy 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAPAKOFLISPPSPVGVGKPISDATPV 130  
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVGKQVEDATPV 126  
Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186  
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186  
Qy 187 RRPGLPP 193  
Db 187 RRPEYTP 193

## RESULT 12

US-09-724-676-61352

; Sequence 61352, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724.676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 61352

; LENGTH: 197

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-61352

Query Match 59.8%; Score 622; DB 5; Length 197;  
Best Local Similarity 63.6%; Pred. No. 1.4e-49;  
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDCVTFQLEKSPRRVRINFSPKSAARARIE 70  
Db 11 SATTACHLDRPFDVGGUCRAKFSLFRTYDKDITFYFKSRVRINFSPFSAADARLQ 70  
Qy 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAPAKOFLISPPSPVGVGKPISDATPV 130  
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPKQFLISPPSPVGVGKQVEDATPV 126  
Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186  
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186  
Qy 187 RRPGLPP 193  
Db 187 RRPEYTP 193

## RESULT 13

US-09-724-676A-61352

; Sequence 61352, Application US/09724676A





```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 63940
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-63940

Query Match      88.7%; Score 923; DB 5; Length 184;
Best Local Similarity 95.6%; Pred. No. 4e-77;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 60

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGKPTIS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGQPIN 120

QY 126 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSPPKPIIQ 185
DB 121 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSPPKPIIQ 180

QY 186 TRR 188
DB 181 GDR 183

RESULT 3
US-09-724-676A-63940
; Sequence 63940, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 63940
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-63940

Query Match      88.7%; Score 923; DB 5; Length 184;
Best Local Similarity 95.6%; Pred. No. 4e-77;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 60

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGKPTIS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGQPIN 120

QY 126 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSPPKPIIQ 185
DB 121 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSPPKPIIQ 180

QY 186 TRR 188
DB 181 GDR 183

US-09-782-953-9.rapn
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; CURRENT FILING DATE: 2002-11-06
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 247500.5
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 11
; OTHER INFORMATION: unknown or other
US-10-290-438-2

Query Match      86.9%; Score 905; DB 6; Length 255;
Best Local Similarity 91.4%; Pred. No. 2.6e-75;
Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 11 STLVCVVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSAARARIE 70
DB 69 NSLFACNVHOSVFEESKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSAARARIE 128

QY 71 LHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGKPIQTRRP 130
DB 129 LHETQFRGKKLYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVGQPIN 188

QY 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSPPKPIQTRRP 190
DB 189 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSPPKPIQTRRP 248

QY 191 LPPSVSN 197
DB 249 LPPSVSN 255

RESULT 5
US-10-290-438-11
; Sequence 11, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; CURRENT FILING DATE: 2002-11-06
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 96017919
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 / Search time 11.6273 Seconds  
(without alignments)  
118.077 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041  
Sequence: 1 MFAPSMDCVSTLVACVVDV.....SPKXIIQTRRGLPPSVSN 197

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 233758 seqs, 65991041 residues

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Pending\_Patents\_AA.New.\*  
1: /cgn2\_6/prodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	980	94.1	192	6	US-10-290-438-10
2	923	88.7	184	5	US-09-724-676-613540
3	923	88.7	184	5	US-09-724-676-613540
4	905	86.9	255	6	US-10-290-438-2
5	658.5	63.3	241	6	US-10-290-438-11
6	641	61.6	197	5	US-09-724-676-61358
7	641	61.6	197	5	US-09-724-676-61359
8	641	61.6	197	5	US-09-724-676-61357
9	641	61.6	197	5	US-09-724-676-61358
10	641	61.6	197	5	US-09-724-676-61359
11	641	61.6	197	5	US-09-724-676-61357
12	622	59.8	197	5	US-09-724-676-61352
13	622	59.8	197	5	US-09-724-676-61352
14	598	57.4	200	5	US-09-724-676-61356
15	598	57.4	200	5	US-09-724-676-61356
16	595	57.2	171	5	US-09-724-676-61354
17	595	57.2	171	5	US-09-724-676-61354
18	482.5	46.3	205	5	US-09-724-676-61360
19	482.5	46.3	205	5	US-09-724-676-61360
20	473.5	45.5	162	5	US-09-724-676-61351
21	473.5	45.5	162	5	US-09-724-676-61351
22	470	45.1	154	5	US-09-724-676-61361
23	470	45.1	154	5	US-09-724-676-61361
24	451	43.3	154	5	US-09-724-676-61353
25	451	43.3	154	5	US-09-724-676-61353
26	427	41.0	157	5	US-09-724-676-61357

27	427	41.0	157	5	US-09-724-676-61357	Sequence 61357, A
28	424	40.7	156	5	US-09-724-676-61355	Sequence 61355, A
29	424	40.7	156	5	US-09-724-676-61355	Sequence 61355, A
30	323	31.0	111	5	US-09-513-999C-5849	Sequence 5849, Ap
31	300	28.8	96	5	US-09-724-676-61363	Sequence 61363, A
32	300	28.8	96	5	US-09-724-676-61363	Sequence 61363, A
33	300	28.8	96	5	US-09-724-676-61363	Sequence 61363, A
34	300	28.8	96	5	US-09-724-676-61363	Sequence 61363, A
35	219	21.0	58	6	US-10-203-138A-12463	Sequence 12463, A
36	174	16.7	56	6	US-10-203-138A-11146	Sequence 11146, A
37	141.5	13.6	104	5	US-09-724-676-61364	Sequence 61364, A
38	141.5	13.6	104	5	US-09-724-676-61364	Sequence 61364, A
39	132.5	12.7	61	5	US-09-724-676-61365	Sequence 61365, A
40	132.5	12.7	61	5	US-09-724-676-61365	Sequence 61365, A
41	89	8.5	722	5	US-09-724-676-90346	Sequence 90346, A
42	89	8.5	722	5	US-09-724-676-90346	Sequence 90346, A
43	89	8.5	724	5	US-09-724-676-90347	Sequence 90347, A
44	89	8.5	724	5	US-09-724-676-90347	Sequence 90347, A
45	89	8.5	807	5	US-09-724-676-90339	Sequence 90339, A

## ALIGNMENTS

RESULT 1  
US-10-290-438-10  
; Sequence 10, Application US/10290438  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; APPLICANT: Streeter, David G.  
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN  
; FILE REFERENCE: PC-0013-1CIP  
; CURRENT APPLICATION NUMBER: US/10/290,438  
; PRIOR FILING DATE: 2002-11-06  
; PRIOR APPLICATION NUMBER: 09/614,474  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PERL Program  
; SEQ ID NO 10  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: g1435040  
US-10-290-438-10

Query Match 94.1%; Score 980; DB 6; Length 192;  
Best Local Similarity 96.4%; Pred. No. 2.6e-82;  
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDDVSTLVACVVDVENVTOYKEKEFEGLFRFYDSCVTFQLEKSPRRVAINFSPKISAA 65  
DB 1 MDDVSTLVACVVDVENVTOYKEKEFEGLFRFYDSCVTFQLEKSPRRVAINFSPKISAA 60

QY 66 RARIELHETOPRGRKGLKLYPAQVOTPETDGDKLHLAPPOKAPFLISPPSPVGMKPIIS 125  
DB 61 RARIELHETOPRGRKGLKLYPAQVOTPETDGDKLHLAPPOKAPFLISPPSPVGMKPIIS 120

QY 126 DATPVNTDYLIVAAKLGPEKYTELHAGTSTSPVVHVCDSDMEEDDPKTSPPKXIIQ 185  
DB 121 DATPVNTDYLIVAAKLGPEKYTELHAGTSTSPVVHVCDSDMEEDDPKTSPPKXIIQ 180

QY 186 TRRRGLPPSVSN 197  
DB 181 TRRRGLPPSVSN 192

RESULT 2  
US-09-724-676-613540  
; Sequence 613540, Application US/09724676

**THIS PAGE BLANK (USPTO)**

COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/853,831  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 2592  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 243 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-853-831-2

Query Match 7.6%; Score 79; DB 3; Length 243;  
Best Local Similarity 25.8%; Pred. No. 0.49;  
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

QY 71 LHETFRGKKLKLYFAOVQTPETDG-----DKLHAPQPAKQFLISPPSSPPVGW 121  
DB 43 LHE-----LYDLDTAPEDNEEAVSQIFPDSVMLA-VQEGIDLTFFPA----- 86  
QY 122 KPISDATPVLYDLVAVAKLGEGKELHAGTESPTSVVHVCD-----SDME 171  
DB 87 -PGSPPEPHLSRQ-----PEQPEQALGVSMPNLVPEVIDLTCHBAGFPSPDDED 136  
QY 172 EEDPKTSPKPK-----IIQTRRPGLPSPV 195  
DB 137 EEGVSEPEPEPEPEPARPTRPKMAPAI 167

RESULT 15  
PCT-US93-09774-2  
Sequence 2, Application PC/TUS9309774  
GENERAL INFORMATION:  
APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION  
TITLE OF INVENTION: METHOD OF INHIBITING REPLICATION OF  
TITILE OF INVENTION: HYPERPROLIFERATIVE CELLS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
CITY: SAN DIEGO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/09774  
FILING DATE: 12-OCT-1992  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: PERKINS, SUSAN M.  
REGISTRATION NUMBER: 36,405  
REFERENCE/DOCKET NUMBER: PP-LJ 9770  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-535-9001

TELEFAX: 619-535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 243 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-09774-2

Query Match 7.6%; Score 79; DB 5; Length 243;  
Best Local Similarity 25.8%; Pred. No. 0.49;  
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

QY 71 LHETFRGKKLKLYFAOVQTPETDG-----DKLHAPQPAKQFLISPPSSPPVGW 121  
DB 43 LHE-----LYDLDTAPEDNEEAVSQIFPDSVMLA-VQEGIDLTFFPA----- 86  
QY 122 KPISDATPVLYDLVAVAKLGEGKELHAGTESPTSVVHVCD-----SDME 171  
DB 87 -PGSPPEPHLSRQ-----PEQPEQALGVSMPNLVPEVIDLTCHBAGFPSPDDED 136  
QY 172 EEDPKTSPKPK-----IIQTRRPGLPSPV 195  
DB 137 EEGVSEPEPEPEPEPARPTRPKMAPAI 167

Search completed: December 11, 2002, 11:39:18  
Job time: 14.2951 secs

```
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPVGV 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKYEELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPEPHLSRQ-----PEQPEQALGPVSMNLPVPEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSKPK-----IIOTRRPGLPPSV 195
Db 137 EEGPVSEPEPEPEPEPARTRRPKMAPAI 167

RESULT 12
US-08-301-316B-2
; Sequence 2, Application US/08301316B
; Patent No. 5776743
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: Method of Sensitizing Tumor Cells with Adenovirus E1A
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/301.316B
; FILING DATE: 06-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-301-316B-2

Query Match 7.6%; Score 79; DB 1; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPVGV 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKYEELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPEPHLSRQ-----PEQPEQALGPVSMNLPVPEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSKPK-----IIOTRRPGLPPSV 195
Db 137 EEGPVSEPEPEPEPEPARTRRPKMAPAI 167

RESULT 13
```

```
US-08-473-399B-2
; Sequence 2, Application US/08473399B
; Patent No. 5866550
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: REVERSE-TRANSFORMATION OF CANCER CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473.399B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,112
; FILING DATE: 13-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1697
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-473-399B-2

Query Match 7.6%; Score 79; DB 2; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPVGV 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKYEELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPEPHLSRQ-----PEQPEQALGPVSMNLPVPEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSKPK-----IIOTRRPGLPPSV 195
Db 137 EEGPVSEPEPEPEPEPARTRRPKMAPAI 167

RESULT 14
US-08-853-831-2
; Sequence 2, Application US/08853831
; Patent No. 6100243
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: Method of Sensitizing Tumor Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
```

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0333 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 559 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 496120  
US-08-884-072-6

Query Match 7.8%; Score 81; DB 2; Length 559;  
Best Local Similarity 25.4%; Pred. No. 0.97; Mismatches 63; Indels 46; Gaps 9;  
Matches 44; Conservative 20;

QY 33 EGLFRTYD--ECVTFQPKSFRVRINFSPKSAARARIELHETQFGKKLKYFAQVOT 90  
DB 22 EGAFKASDQEMTFRFQHLHEV--GYAAPSLPQTR-----RLRV----- 61  
QY 91 PETDGDKHLAPQAPKQFLISPPSSPPVGKPISDATPYLAND---LLYAVAKLGPG 146  
DB 62 -DHSVTSLSH-DPLFEEQREVPQSSP-----EDIPYEEBWPFLPNVDKAGPAV 111  
QY 147 KYELHAGTSTSPVVVHVCDSDMEEDPKTSKPKTIQ-----TRRPLGP 193  
DB 112 PQBAIPLOKQEPPOVHI-----EQKEIDPPAQPOEELVQKEVPHLAGQLPP 160

RESULT 10  
US-09-212-168-6  
Sequence 6, Application US/09212168  
Patent No. 6303765  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PaeuseQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/212,168  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/884,072  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 559 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 496120  
US-09-212-168-6

Query Match 7.8%; Score 81; DB 4; Length 559;  
Best Local Similarity 25.4%; Pred. No. 0.97; Mismatches 63; Indels 46; Gaps 9;  
Matches 44; Conservative 20;

QY 33 EGLFRTYD--ECVTFQPKSFRVRINFSPKSAARARIELHETQFGKKLKYFAQVOT 90  
DB 22 EGAFKASDQEMTFRFQHLHEV--GYAAPSLPQTR-----RLRV----- 61  
QY 91 PETDGDKHLAPQAPKQFLISPPSSPPVGKPISDATPYLAND---LLYAVAKLGPG 146  
DB 62 -DHSVTSLSH-DPLFEEQREVPQSSP-----EDIPYEEBWPFLPNVDKAGPAV 111  
QY 147 KYELHAGTSTSPVVVHVCDSDMEEDPKTSKPKTIQ-----TRRPLGP 193  
DB 112 PQBAIPLOKQEPPOVHI-----EQKEIDPPAQPOEELVQKEVPHLAGQLPP 160

RESULT 11  
US-07-960-112B-2  
Sequence 2, Application US/07960112B  
Patent No. 5516631  
GENERAL INFORMATION:  
APPLICANT: Frisch, Steven M.  
TITLE OF INVENTION: Method of Inhibiting Replication of  
HYPERPROLIFERATIVE CELLS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/960,112B  
FILING DATE: 13-OCT-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9429  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 243 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-960-112B-2

Query Match 7.6%; Score 79; DB 1; Length 243;



City: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/574,959A  
FILING DATE: 19-DEC-95  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: DFN-008  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 905 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-574-959A-9

Query Match  
Best Local Similarity 29.6%; Score 89; DB 2; Length 905;  
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160  
DB 751 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 801

QY 161 VHAVCDSDME---EEDPKTSPKPKIITQTRRPLGPSV 195  
DB 802 APEALPSQGEVEREGSPAGPPQELVEEPEXPFTL 839

RESULT 5  
US-09-357-014-9  
Sequence 9, Application US/09357014  
Patent No. 6291645  
GENERAL INFORMATION:  
APPLICANT: Jaekyoon Shin, Insil Jouny, Ratna K. Vadlamudi  
and Jack L. Strominger  
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES  
AND USES THEREFOR  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/357,014  
FILING DATE: 19-JUL-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/574,959  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: DFN-008  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 905 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-09-357-014-9

Query Match  
Best Local Similarity 29.6%; Score 89; DB 4; Length 905;  
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160  
DB 751 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 801

QY 161 VHAVCDSDME---EEDPKTSPKPKIITQTRRPLGPSV 195  
DB 802 APEALPSQGEVEREGSPAGPPQELVEEPEXPFTL 839

RESULT 6  
US-08-574-959A-7  
Sequence 7, Application US/08574959A  
Patent No. 5562224  
GENERAL INFORMATION:  
APPLICANT: Jaekyoon Shin, Insil Jouny, Ratna K. Vadlamudi  
and Jack L. Strominger  
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES  
AND USES THEREFOR  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/574,959A  
FILING DATE: 19-DEC-95  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: DFN-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1135 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-574-959A-7

Query Match  
Best Local Similarity 29.6%; Score 89; DB 2; Length 1135;  
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160  
DB 981 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 1031

QY 161 VHAVCDSDME---EEDPKTSPKPKIITQTRRPLGPSV 195





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.2951 Seconds

(without alignments)  
513.170 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACWV.....SPKXIIQTRRPGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:\*

2: /cgn2\_6/prodata/1/aa/5A.COMB.pep:\*

3: /cgn2\_6/prodata/1/aa/5B.COMB.pep:\*

4: /cgn2\_6/prodata/1/aa/6A.COMB.pep:\*

5: /cgn2\_6/prodata/1/aa/6B.COMB.pep:\*

6: /cgn2\_6/prodata/1/aa/6C.COMB.pep:\*

7: /cgn2\_6/prodata/1/aa/6D.COMB.pep:\*

8: /cgn2\_6/prodata/1/aa/6E.COMB.pep:\*

9: /cgn2\_6/prodata/1/aa/6F.COMB.pep:\*

10: /cgn2\_6/prodata/1/aa/6G.COMB.pep:\*

11: /cgn2\_6/prodata/1/aa/6H.COMB.pep:\*

12: /cgn2\_6/prodata/1/aa/6I.COMB.pep:\*

13: /cgn2\_6/prodata/1/aa/6J.COMB.pep:\*

14: /cgn2\_6/prodata/1/aa/6K.COMB.pep:\*

15: /cgn2\_6/prodata/1/aa/6L.COMB.pep:\*

16: /cgn2\_6/prodata/1/aa/6M.COMB.pep:\*

17: /cgn2\_6/prodata/1/aa/6N.COMB.pep:\*

18: /cgn2\_6/prodata/1/aa/6O.COMB.pep:\*

19: /cgn2\_6/prodata/1/aa/6P.COMB.pep:\*

20: /cgn2\_6/prodata/1/aa/6Q.COMB.pep:\*

21: /cgn2\_6/prodata/1/aa/6R.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	586	56.3	171	2	US-08-665-040-2
2	101	9.7	454	3	US-08-348-518C-4
3	101	9.7	454	3	US-08-476-509B-4
4	89	8.5	905	2	US-08-574-959A-9
5	89	8.5	905	2	US-09-357-014-9
6	89	8.5	1135	2	US-08-574-959A-7
7	89	8.5	1135	2	US-09-357-014-7
8	81	7.8	434	2	US-08-815-718-3
9	81	7.8	559	2	US-08-884-072-6
10	81	7.8	559	2	US-09-212-168-6
11	79	7.6	243	1	US-07-560-112B-2
12	79	7.6	243	1	US-08-301-316B-2
13	79	7.6	243	2	US-08-473-399B-2
14	79	7.6	243	2	US-08-853-831-2
15	79	7.6	243	5	PCT-US93-09774-2
16	78	7.5	1018	1	US-08-408-093-6
17	78	7.5	1018	1	US-08-408-420A-6
18	78	7.5	1018	1	US-08-714-901-6
19	78	7.5	1018	1	US-08-040-741-6
20	76.5	7.3	330	4	US-09-145-391-2
21	76	7.3	1958	1	US-07-945-283-2
22	75.5	7.3	154	4	US-08-866-928B-9
23	75	7.2	401	2	US-08-549-004A-5
24	75	7.2	401	2	US-09-051-982A-5
25	75	7.2	976	3	US-08-560-005-2
26	75	7.2	976	3	US-09-195-868-14
27	75	7.2	976	4	US-09-418-540-2

28	75	7.2	1187	4	US-08-664-962B-8	Sequence 8, Appl
29	75	7.2	1187	4	US-09-311-743-8	Sequence 8, Appl
30	75	7.2	1187	4	US-09-195-868-15	Sequence 15, Appl
31	75	7.2	1229	3	US-09-195-868-28	Sequence 28, Appl
32	75	7.2	2441	1	US-08-194-468-2	Sequence 2, Appl
33	75	7.2	2441	3	US-08-961-739-2	Sequence 2, Appl
34	75	7.2	2441	4	US-09-514-247A-8	Sequence 8, Appl
35	72.5	7.0	659	4	US-09-562-737-12	Sequence 12, Appl
36	71	6.8	350	2	US-08-960-023-8	Sequence 8, Appl
37	71	6.8	357	1	US-08-145-006C-12	Sequence 12, Appl
38	71	6.8	357	5	PCT-US94-00945-12	Sequence 12, Appl
39	71	6.8	2972	4	US-09-579-181-2	Sequence 12, Appl
40	71	6.8	3118	4	US-09-579-181-1	Sequence 1, Appl
41	70.5	6.8	249	1	US-08-450-246-57	Sequence 57, Appl
42	70.5	6.8	249	1	US-08-450-246-57	Sequence 57, Appl
43	70.5	6.8	249	1	US-08-450-098-57	Sequence 57, Appl
44	70.5	6.8	249	1	US-08-451-235-57	Sequence 57, Appl
45	70.5	6.8	249	1	US-08-450-236-57	Sequence 57, Appl

## ALIGNMENTS

## RESULT 1

US-08-665-040-2  
Sequence 2, Application US/08665040

Patent No. 5869318

GENERAL INFORMATION:

APPLICANT: ESTIVILL, PALEJA, XAVIER

APPLICANT: FUENTES, JUAN JOSE

APPLICANT: PRITCHARD, MELANIE

TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE

TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,

TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING PCR" TECHNIQUE

TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (OSCR1) HIGHLY

TITLE OF INVENTION: EXPRESSED IN FORTLE BRAIN AND IN HEART AND METHOD

TITLE OF INVENTION: FOR CHARACTERIZING IT.

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: LADAS & PARRY

STREET: 26 WEST 61ST STREET

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORD PERFECT 5.1 FOR DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/665,040

FILING DATE: JUNE 7, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: ES P9501140

FILING DATE: JUNE 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: JANET I. CORD

REGISTRATION NUMBER: 33,778

REFERENCE/DOCKET NUMBER: U010815-9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 708-1800

TELEFAX: (212) 246-8959

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 171 amino acids

TYPE: amino acid

STRANDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-665-040-2

Qy 110 ProAlaSerProProValGlyThrPheGlnValGluAspAlaThrProValIleAsnTyr 129  
Db 335 CCCGCCCTCCGCCAGTGGGATGAACAGTGAAGATGCCGCCAGTCTATAAATAT 394  
Qy 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149  
Db 395 GATCTTTATATGTCATCTCAAGCTGGGCCAGGGGAAAGATGATGAATTCACGCGC 454  
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
Db 455 ACTGACACCACTCCAGCGCTGGTGGTCCATGTATGTGAGAGTATCAAGAG- --AAGGAG 511  
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysPheProLysLysLysLysLysLys 189  
Db 512 GAAGAAGAGGAATGGAAGAATGAGGAGACCTAAGCCAAAATTTATCCAGACCGAGG 571  
Qy 190 ProGluTyrThrProIleHisLeuSer 198  
Db 572 CCGAGTACACGCCGATCCACTCAGC 598  
RESULT 14  
US-09-724-676-12742  
; Sequence 12742, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12742  
; LENGTH: 1056  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676-12742  
Alignment Scores:  
Pred. No.: 1,04e-72 Length: 1056  
Score: 731.00 Matches: 138  
Percent Similarity: 97.24% Conservatives: 3  
Best Local Similarity: 95.17% Mismatches: 4  
Query Match: 69.75% Indels: 0  
Gaps: 0  
DB:  
US-09-782-953-3 (1-198) x US-09-724-676-12742 (1-1056)  
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCCCGCCAGCCACCTGCTGTACCTGGACCCG 84  
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
Db 85 CCGCTGTTCTGGAGCGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144  
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
Db 145 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGATAAATTCAGCAAC 204  
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 205 CCCCTTCGCCAGCAGATGCCAGCTCCAGCTGCATAAGACTGAGTTCTCGGAAGGAA 264  
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProAsn 100  
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324  
Qy 101 ProAspLysGlnPheLeuIleSerProAlaSerProValGlyThrLysGlnVal 120  
Db 325 CCAGCAAGCAGAGTTCCTGATCTCCCTCCCTCCCGCAGTGGGATGGAACAAAGT 384  
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 385 GAAGATGCGACCCAGTCATTAACCTATGATCTCTTATATGCACTCTCCAAAGTGGGCA 444  
Qy 141 GlyGluLysTyrGlu 145  
Db 445 GGTGATTGGCTAGAG 459  
Search completed: December 15, 2002, 00:04:20  
Job time : 57.7555 secs

Db 385 GAAGATGCGACCCAGTCATAAATACTATGATCTCTTATATGCACTCTCCAAAGTGGGCA 444  
Qy 141 GlyGluLysTyrGlu 145  
Db 445 GGTGATTGGCTAGAG 459  
RESULT 15  
US-09-724-676A-12742  
; Sequence 12742, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12742  
; LENGTH: 1056  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-724-676A-12742  
Alignment Scores:  
Pred. No.: 1,04e-72 Length: 1056  
Score: 731.00 Matches: 138  
Percent Similarity: 97.24% Conservatives: 3  
Best Local Similarity: 95.17% Mismatches: 4  
Query Match: 69.75% Indels: 0  
Gaps: 0  
DB:  
US-09-782-953-3 (1-198) x US-09-724-676A-12742 (1-1056)  
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCCCGCCAGCCACCTGCTGTACCTGGACCCG 84  
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40  
Db 85 CCGCTGTTCTGGAGCGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144  
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60  
Db 145 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGATAAATTCAGCAAC 204  
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80  
Db 205 CCCCTTCGCCAGCAGATGCCAGCTCCAGCTGCATAAGACTGAGTTCTCGGAAGGAA 264  
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProAsn 100  
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324  
Qy 101 ProAspLysGlnPheLeuIleSerProAlaSerProValGlyThrLysGlnVal 120  
Db 325 CCAGCAAGCAGAGTTCCTGATCTCCCTCCCTCCCGCAGTGGGATGGAACAAAGT 384  
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140  
Db 385 GAAGATGCGACCCAGTCATTAACCTATGATCTCTTATATGCACTCTCCAAAGTGGGCA 444  
Qy 141 GlyGluLysTyrGlu 145  
Db 445 GGTGATTGGCTAGAG 459  
Search completed: December 15, 2002, 00:04:20  
Job time : 57.7555 secs

```

DB: 5 Gaps: 1
US-09-782-953-3 (1-198) x US-09-724-676A-12743 (1-1839)
QY 30 AAlaLysPheGluSerLeuPheArgThrTyraPlyAspThrThrPheGlnTyrrPheLys 49
   |||
Db 58 GCCAAATTGAGTCCCTCTTTAGACGTAATGACAGACATCACCTTTCAGTATTTTAAG 117
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLysSerAlaAlaAspAlaArgLeu 69
   |||
Db 118 AGCTTCAACAGCTCAGAAATTAATTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeu 89
   |||
Db 178 CAGCTGCATTAAGACTGAGTTCTTGGAAGAGAAATGAAGTATATTTTGCTCACAACCTTA 237
QY 90 HisIleGlySerSerHisLysLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109
   |||
Db 238 CACATAGGAAGCTCACACCTGCTCCGCCAAATCCAGACAGACGTTTCTGATCTCCCT 297
QY 110 ProAlaSerProProValGlyTyrrPlyGlnValGluAspAlaThrProValIleAsnTyrr 129
   |||
Db 298 CCCGCTCTCCGCCAGTGGAGTGAAGCAAGTGAAGTGGACCCAGATCATTAACCTAT 357
QY 130 AspLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAla 149
   |||
Db 358 GATCTCTTAATATGCATCTCCAACTGGGCGCAGGGGAAAAGTGAATTGCACGACGCG 417
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
   |||
Db 418 ACTGACACCACTCCCACTGGTGGTGTCCATGTATGTAGTAGTCAAGAG--AAGGAG 474
QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
   |||
Db 475 GAAGAAAGAGAAATGAAAGATGAGAGACCTTAAGCCAAATAATTCACGACGAGAGG 534
QY 190 ProGluTyrrThrProIleHisLysSer 198
   |||
Db 535 CCGAGTACACGCCGATCCACTCAGC 561

RESULT 12
US-09-724-676-12745
; Sequence 12745, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12745
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12745

Alignment Scores:
Pred. No.: 2,196-85 Length: 1876
Score: 847.50 Matches: 162
Percent Similarity: 97.63% Conservative: 3
Best Local Similarity: 95.86% Mismatches: 3
Query Match: 80.87% Indels: 1
Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12745 (1-1876)
QY 30 AAlaLysPheGluSerLeuPheArgThrTyraPlyAspThrThrPheGlnTyrrPheLys 49
   |||
Db 95 GCCAAATTGAGTCCCTCTTTAGACGTAATGACAGACATCACCTTTCAGTATTTTAAG 154
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLysSerAlaAlaAspAlaArgLeu 69
   |||

```

```

Db 155 AGCTTCAACAGCTCAGAAATTAATTCAGCAACCCCTTCTCCGACGACAGTCCAGGCTC 214
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeu 89
   |||
Db 215 CACCTGCATTAAGCTGAGTTCTGGGAAGAAATGAAGTATATTTTGCTCAGACCTTA 274
QY 90 HisIleGlySerSerHisLysLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109
   |||
Db 275 CACATAGGAAGCTCACACCTGCTCCGCCAAATCCAGACAGACGTTTCTGATCTCCCT 334
QY 110 ProAlaSerProProValGlyTyrrPlyGlnValGluAspAlaThrProValIleAsnTyrr 129
   |||
Db 335 CCCGCTCTCCGCCAGTGGAGTGAAGCAAGTGAAGATGCGACCCAGTCAATTAACCTAT 394
QY 130 AspLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAla 149
   |||
Db 395 GATCTCTTAATATCCATCTCCAACTGGGGCCAGGGGAAAAGTGAATTGCACGCGCGC 454
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
   |||
Db 455 ACTGACACCACTCCCACTGGTGGTGTCCATGTATGTAGAGTGAATCAAG--AAGGAG 511
QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
   |||
Db 512 GAAGAAAGAGAAATGAAAGATGAGAGACCTTAAGCCAAATAATTCACGACGAGAGG 571
QY 190 ProGluTyrrThrProIleHisLysSer 198
   |||
Db 572 CCGAGTACACGCCGATCCACTCAGC 598

RESULT 13
US-09-724-676A-12745
; Sequence 12745, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12745
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12745

Alignment Scores:
Pred. No.: 2,196-85 Length: 1876
Score: 847.50 Matches: 162
Percent Similarity: 97.63% Conservative: 3
Best Local Similarity: 95.86% Mismatches: 3
Query Match: 80.87% Indels: 1
Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12745 (1-1876)
QY 30 AAlaLysPheGluSerLeuPheArgThrTyraPlyAspThrThrPheGlnTyrrPheLys 49
   |||
Db 95 GCCAAATTGAGTCCCTCTTTAGACGTAATGACAGACATCACCTTTCAGTATTTTAAG 154
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLysSerAlaAlaAspAlaArgLeu 69
   |||
Db 155 AGCTTCAACAGCTCAGAAATTAATTCAGCAACCCCTTCTCCGACGACAGTCCAGGCTC 214
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeu 89
   |||
Db 215 CAGCTGCATTAAGCTGAGTTCTGGGAAGAAAGAAAGTATATTTTGCTCAGACCTTA 274
QY 90 HisIleGlySerSerHisLysLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109
   |||
Db 275 CACATAGGAAGCTCACACCTGCTCCGCCAAATCCAGACAGACGTTTCTGATCTCCCT 334

```

```
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 2348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 042176.5
US-10-240-965-60

Alignment Scores:
Pred. No.: 1.35e-86      Length: 2348
Score: 859.50           Matches: 168
Percent Similarity: 92.59%      Conservative: 7
Best Local Similarity: 88.89%    Mismatches: 12
Query Match: 82.01%            Indels: 2
DB: 6                      Gaps: 1

US-09-782-953-3 (1-198) x US-10-240-965-60 (1-2348)
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 181 AGCTCCCTGATTGCTGTGTGGCAACAGTGTATCTTTCAGGAAAGTGAACACGAGGCC 240
Qy 31 LysPheGluSerLeuPheArgThrTyraAspLysAspThrThrPheGlnTyPheLysSer 50
Db 241 AAATTTGAGTCCCTCTTTAGGACGTATGACCAAGGACATCACCTTTTCAGTATTTTAAAGAGC 300
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 301 TTCAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 360
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeuHis 90
Db 361 CTGCATAGACTGATTTCTCGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTACAC 420
Qy 91 IleGlySerSerHisLeu-AlaProAsnProAspLysGlnPheLeuIleSerProPr 110
Db 421 ATAGGAAGCTCACACCTCGGCTCGGCCAAATCCAGCAAGCAGTTTCTGTATCTCCCTCC 480
Qy 110 oAlaSerProProValGlyTbLysGlnValGluAspAlaThrProValIleAsnTyCas 130
Db 481 CGCTCTCCGCGAGTGGATGGAACAAGTGAAGATCGACCCCGCATATAAATCTATGA 540
Qy 130 pLeuLeuTyraIleSerLysLeuGlyProGlyLysTyPheGluLeuHisAlaAlaTh 150
Db 541 TCTCTTATGCTATCTCCAAAGCTGGGCGCCAGGGGAAAGTATGAATTTGCACGCGAGC 600
Qy 150 rAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGl 170
Db 601 TGACACCACTCCACGCGTGGTGGTCCATGTATGTGAGAGTGTATCAAGAG---AAGGAGGA 657
Qy 170 uGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPr 190
Db 658 AGAAGAGGAATGGAAGAATGAGGAGACCTTAGCCAAAATTTATCCAGACGAGGAGGCC 717
Qy 190 oGluTyThrProIleHisLeuSer 198
Db 718 GGAGTACACGCGCATCCACCTCAGC 742

RESULT 10
US-09-724-676-12743
; Sequence 12743, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12743

Alignment Scores:
Pred. No.: 2.13e-85      Length: 1839
Score: 847.50           Matches: 162
Percent Similarity: 97.63%      Conservative: 3
Best Local Similarity: 95.86%    Mismatches: 3
Query Match: 80.87%            Indels: 1
DB: 5                      Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12743 (1-1839)
Qy 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrThrPheGlnTyPheLys 49
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAAG 117
Qy 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 118 AGCTTCAAACGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
Qy 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeu 89
Db 178 CAGCTGCATGAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACCTTA 237
Qy 90 HisIleGlySerSerHisLeuAlaProAsnProAsnProAspLysGlnPheLeuIleSerPro 109
Db 238 CACATAGAGAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTCGTATCTCCCT 297
Qy 110 ProLaserProProValGlyTbLysGlnValGluAspAlaThrProValIleAsnTy 129
Db 298 CCCGCCCTCTCCGCGAGTGGATGGAACAAGTGAAGATCGACCCCGCATATAAATCTAT 357
Qy 130 AspLeuLeuTyraIleSerLysLeuGlyProGlyLysTyPheGluLeuHisAlaAla 149
Db 358 GATCTCTTATATGCTATCTCCAAAGCTGGGCGCCAGGGGAAAGTATGAATTTGCACGCGAGC 417
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 418 ACTGACACCACTCCACGCGTGGTGGTCCATGTATGTGAGAGTGTATCAAGAG---AAGGAG 474
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
Db 475 GAAGAAGAGGAATGGAAGAATGAGGAGACCTAAGCCAAAATTTATCCAGACGAGGAGG 534
Qy 190 ProGluTyThrProIleHisLeuSer 198
Db 535 CCGGAGTACACGCGCATCCACCTCAGC 561

RESULT 11
US-09-724-676A-12743
; Sequence 12743, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12743

Alignment Scores:
Pred. No.: 2.13e-85      Length: 1839
Score: 847.50           Matches: 162
Percent Similarity: 97.63%      Conservative: 3
Best Local Similarity: 95.86%    Mismatches: 3
Query Match: 80.87%            Indels: 1
DB: 5                      Gaps: 1
```

Score: 872.50 Matches: 168  
Percent Similarity: 93.09% Conservative: 7  
Best Local Similarity: 89.36% Mismatches: 12  
Query Match: 83.25% Indels: 1  
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12747 (1-2297)

```
OY 11 SerAlaThrIleAlaCyHisLeuAspProArgValPheValAspGlyLeuCyArgAla 30
DB 93 AGCTCCCTGATTGCTGTGGCAACAGTATCTTCAGCGCAAGTGAACAGGAGGCC 152
OY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
DB 153 AAATTGAGTCCCTCTTAGACGTATGACAGACATCACTTTCAGTATTTTAAAGC 212
OY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70
DB 213 TTCAACGAGTCAGATTAACCTTACGAAACCCCTTCTCCGACGATGCCAGGCTCCAG 272
OY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
DB 273 CTCGATTAAGCTAGATTTCGGGAAAGAAATGAAGTTATATTGCTCAGACCTTACAC 332
OY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
DB 333 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAAAGAGTTTCGATCTCCCTCCC 392
OY 111 AlSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
DB 393 GCCTCTCCGACGAGGATGGAACAAAGTGAAGATGCGACCCAGCATTAACCTATGAT 452
OY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
DB 453 CTCCTTATATCCATCTCCAAAGCTGGGCCACGGGAAAGTATGATTCACGCGACACT 512
OY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
DB 513 GACACCACTCCACGCGTGGTGCATGTATGTAGAGTATCAAGAG--AAGAGAGAA 569
OY 171 GluGlnGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
DB 570 GAAGAGAAATGAAAGAAATGAGAGACTTAAGCCAAATTAATCCAGACCGAGGCGC 629
OY 191 GluTyrThrProIleHisLeuSer 198
DB 630 GAGTACAGCGCGATCCACCTCAGC 653
```

RESULT 8

```
US-09-724-676A-12747
; Sequence 12747, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
```

LOCATION: (23)..(23)  
OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12747

Alignment Scores:

Prod. No.: 4,63e-88 Length: 2297  
Score: 872.50 Matches: 168  
Percent Similarity: 93.09% Conservative: 7  
Best Local Similarity: 89.36% Mismatches: 12  
Query Match: 83.25% Indels: 1  
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12747 (1-2297)

```
OY 11 SerAlaThrIleAlaCyHisLeuAspProArgValPheValAspGlyLeuCyArgAla 30
DB 93 AGCTCCCTGATTGCTGTGGCAACAGTATCTTCAGCGCAAGTGAACAGGAGGCC 152
OY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
DB 153 AAATTGAGTCCCTCTTAGACGTATGACAGACATCACTTTCAGTATTTTAAAGC 212
OY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70
DB 213 TTCAACGAGTCAGATTAACCTTACGAAACCCCTTCTCCGACGATGCCAGGCTCCAG 272
OY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
DB 273 CTCGATTAAGCTAGATTTCGGGAAAGAAATGAAGTTATATTGCTCAGACCTTACAC 332
OY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
DB 333 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAAAGAGTTTCGATCTCCCTCCC 392
OY 111 AlSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
DB 393 GCCTCTCCGACGAGGATGGAACAAAGTGAAGATGCGACCCAGCATTAACCTATGAT 452
OY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
DB 453 CTCCTTATATCCATCTCCAAAGCTGGGCCACGGGAAAGTATGATTCACGCGACACT 512
OY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
DB 513 GACACCACTCCACGCGTGGTGCATGTATGTAGAGTATCAAGAG--AAGAGAGAA 569
OY 171 GluGlnGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
DB 570 GAAGAGAAATGAAAGAAATGAGAGACTTAAGCCAAATTAATCCAGACCGAGGCGC 629
OY 191 GluTyrThrProIleHisLeuSer 198
DB 630 GAGTACAGCGCGATCCACCTCAGC 653
```

RESULT 9

```
US-10-240-965-60
; Sequence 60, Application US/10240965
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, DOV
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAMN, Richard M.
; APPLICANT: SELHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
```

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
Db 212 AGCTCCCTGATTGCTGTGGCAACAGTAGTATCTTCACGGAAGTGAACACGAGGCC 271  
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTCAGTATTTTAAGAGC 331  
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391  
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
Db 392 CTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGCTCAGACCTTTACAC 451  
Qy 91 IleGlySerSerHisLeuAlaProProAsnProAsnProAspLysGlnPheLeuLysSerProPro 110  
Db 452 ATAGGAAGCTCACCTCGCTCCGCAATCCAGAAAGCAGTTTCTGATCTCCCTCC 511  
Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
Db 512 GCCTCTCCGCGAGTGGGATGGAACAGTGAAGATGCGACCCAGTCATAAATCTATGAT 571  
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
Db 572 CTCTTATATGCTCATCTCCAAGCTGGGCGCAGGGGAAAAGTATGAATTCACGCGAGCT 631  
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170  
Db 632 GACACACTCCCGAGTGGTGTCTGATGTGAGAGTGATCAAGAG---AAGGAGGAA 688  
Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190  
Db 689 GAAGAGGAATGGAAGATGAGGAGACCTTAAGCCAAAATTTATCCAGACGAGGAGCGC 748  
Qy 191 GluTyrThrProIleHisLeuSer 198  
Db 749 GAGTACAGCGCGATCCACCTCAGC 772

RESULT 6  
US-09-724-676A-12756  
; Sequence 12756, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12756  
; LENGTH: 2050  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (47)..(47)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676A-12756

Alignment Scores:  
Pred. No.: 3.97e-88 Length: 2050  
Score: 872.50 Matches: 168  
Percent Similarity: 93.09% Conservative: 7  
Best Local Similarity: 89.36% Mismatches: 12  
Query Match: 83.25% Indels: 1  
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12756 (1-2050)  
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30  
Db 212 AGCTCCCTGATTGCTGTGGCAACAGTAGTATCTTCACGGAAGTGAACACGAGGCC 271  
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50  
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTCAGTATTTTAAGAGC 331  
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70  
Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391  
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90  
Db 392 CTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGCTCAGACCTTTACAC 451  
Qy 91 IleGlySerSerHisLeuAlaProProAsnProAsnProAspLysGlnPheLeuLysSerProPro 110  
Db 452 ATAGGAAGCTCACCTCGCTCCGCAATCCAGAAAGCAGTTTCTGATCTCCCTCC 511  
Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130  
Db 512 GCCTCTCCGCGAGTGGGATGGAACAGTGAAGATGCGACCCAGTCATAAATCTATGAT 571  
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150  
Db 572 CTCTTATATGCTCATCTCCAAGCTGGGCGCAGGGGAAAAGTATGAATTCACGCGAGCT 631  
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170  
Db 632 GACACACTCCCGAGTGGTGTCTGATGTGAGAGTGATCAAGAG---AAGGAGGAA 688  
Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190  
Db 689 GAAGAGGAATGGAAGATGAGGAGACCTTAAGCCAAAATTTATCCAGACGAGGAGCGC 748  
Qy 191 GluTyrThrProIleHisLeuSer 198  
Db 749 GAGTACAGCGCGATCCACCTCAGC 772

RESULT 7  
US-09-724-676-12747  
; Sequence 12747, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12747  
; LENGTH: 2297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (18)..(18)  
; OTHER INFORMATION: n is a,c,g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (23)..(23)  
; OTHER INFORMATION: n is a,c,g, or t  
US-09-724-676-12747  
Alignment Scores:  
Pred. No.: 4.63e-88 Length: 2297

```
OY 51 PheLYsArGVaIARgIIeaSnPheSerAnProLeuSerAlaIAspAlaArgLeuArg 70
Db 213 TTCAAACGAGTAgATAACTTCAGCAACCTTCTCCGACGAGATCCAGGCTCCAG 272
OY 71 LeuHISLyThrgIuPheLeuGlyLygIuMeLyLeuTyRPhaIaGIInThreLHis 90
Db 273 CTGCATTAAGACTGAGATTTCTGGGAAAGAAATGAAGTTATTTTCTCAGACTTACAC 332
OY 91 ILeGlySerSerHISLeuAlaProAnProAplygInPheLeuIleSerProPro 110
Db 333 ATAGCAACCTCACACTGCTCCGCCAATCCAGACAGAGTTTCTGATCTCCCTCC 392
OY 111 AlAserProProValGlyTyRlygInValGluAspAlaThrProValIleAsnTyRAsp 130
Db 393 GCCTCTCCGCACTGGGAGTGAAGAAACAGTGAAGATGGACCCCACTATTAACCTATGAT 452
OY 131 LeuLeuTyRAlaIleSerLyLeuGlyProGlyGlyTyRGlulLeuHISAlaIaThr 150
Db 453 CTCTTATATGCAATCTCCAGCTGGGCGCAGGGAAGTATGAATTCACGCGACCT 512
OY 151 AspProThrProSerValValValHISValCysGlySerAspGInGluAsnGluGlu 170
Db 513 GACACCACTCCAGCGTGTGTCTCATGTATGTAGAGTATCAAGAG--AAGGAGAA 569
OY 171 GluGluGluMeGluArgMeLyAsrProLyProLyIleIleGInThraArgPro 190
Db 570 GAAGAGGAATGGAAGAAATGAAGACCTTAAGCCAAATATTCAGACCAAGAGCGCG 629
OY 191 GluTyRThrProIleHISLeuSer 198
Db 630 GAGTACAGCGCCGATCCACTCAGC 653

RESULT 4
US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12748

Alignment Scores:
Pred. No.: 3,66e-88 Length: 1931
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12748 (1-1931)
OY 11 SeArIaThrIleAlaCyHISLeuAspProArgValPheValaAspGlyLeuCyArgAla 30
||||: ||||||| :||| :|||
```

```
Db 93 AGCTCCCTGATTCCTGTGTGGCAACAGTATATCTTCAGCCAAAGTAAACAGGCGC 152
OY 31 LyPheGluSerLeuPheArgThrTyRAspLyAspThrPheGInTyRPhelySer 50
Db 153 AAATTTAGTCCCTCTTTAGACGTATGACAGACATCCCTTTCAGTATTTTAAGAGC 212
OY 51 PheLYsArGVaIARgIIeaSnPheSerAnProLeuSerAlaIAspAlaArgLeuArg 70
Db 213 TTCAAACGAGTAgATAACTTCAGCAACCTTCTCCGACGAGATCCAGGCTCCAG 272
OY 71 LeuHISLyThrgIuPheLeuGlyLygIuMeLyLeuTyRPhaIaGIInThreLHis 90
Db 273 CTGCATTAAGACTGAGATTTCTGGGAAAGAAATGAAGTTATTTTCTCAGACTTACAC 332
OY 91 ILeGlySerSerHISLeuAlaProAnProAplygInPheLeuIleSerProPro 110
Db 333 ATAGCAACCTCACACTGCTCCGCCAATCCAGACAGAGTTTCTGATCTCCCTCC 392
OY 111 AlAserProProValGlyTyRlygInValGluAspAlaThrProValIleAsnTyRAsp 130
Db 393 GCCTCTCCGCACTGGGAGTGAAGAAACAGTGAAGATGGACCCCACTATTAACCTATGAT 452
OY 131 LeuLeuTyRAlaIleSerLyLeuGlyProGlyGlyTyRGlulLeuHISAlaIaThr 150
Db 453 CTCTTATATGCAATCTCCAGCTGGGCGCAGGGAAGTATGAATTCACGCGACCT 512
OY 151 AspProThrProSerValValValHISValCysGlySerAspGInGluAsnGluGlu 170
Db 513 GACACCACTCCAGCGTGTGTCTCATGTATGTAGAGTATCAAGAG--AAGGAGAA 569
OY 171 GluGluGluMeGluArgMeLyAsrProLyProLyIleIleGInThraArgPro 190
Db 570 GAAGAGGAATGGAAGAAATGAAGACCTTAAGCCAAATATTCAGACCAAGAGCGCG 629
OY 191 GluTyRThrProIleHISLeuSer 198
Db 630 GAGTACAGCGCCGATCCACTCAGC 653

RESULT 5
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756

Alignment Scores:
Pred. No.: 3,97e-88 Length: 2050
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12756 (1-2050)
```



```
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValAlaGlnLeuAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTTCAAGAGCTTCAACAGAGTCAGAATAAACTTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCTTCTCCGAGCAGATGCCAGCTCAGCTGCATAGACTGAGTTCTGGGAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrPlyGlnVal 120
Db 325 CCAGAACAGAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAACAAGTG 384
Qy 121 GluAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaLeuSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCCGCTCATAACTATGATCTCTTATATGCTCTCCAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAGATGATGAATTCACGAGCGACTGACACCACTCCAGCGTGGTCCATGTA 504
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTATCAAGAG---AAGGAGGAAGAGGAATGGAAAGAAATGAGGAGACCT 561
Qy 181 LysProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
Db 562 AAGCCAAAATTTATCCAGACGAGGAGCGGAGTACACGCCGATCCACCTCAGC 615

RESULT 2
US-09-724-676A-12741
; Sequence 12741, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741

Alignment Scores:
Pred. No.: 2,29e-102 Length: 1893
Score: 999.50 Matches: 190
Percent Similarity: 97.98% Conservative: 4
Best Local Similarity: 95.96% Mismatches: 3
Query Match: 95.37% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12741 (1-1893)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCCCGAGCCACCATCGCTGTCCACCTGACCCG 84
Qy 21 ArgValPheValAspGlyLysCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CCGGTGTTCTGGACCGCTGTGCGGGCCAATTTAGTCCCTCTTTAGGACGATGAC 144
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValAlaGlnLeuAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTTCAAGAGCTTCAACAGAGTCAGAATAAACTTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCTTCTCCGAGCAGATGCCAGCTCAGCTGCATAGACTGAGTTCTGGGAAGGAA 264
```

```
Db 205 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTCTGGGAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrPlyGlnVal 120
Db 325 CCAGAACAGAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAACAAGTG 384
Qy 121 GluAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaLeuSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCCGCTCATAACTATGATCTCTTATATGCTCTCCAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAGATGATGAATTCACGAGCGACTGACACCACTCCAGCGTGGTCCATGTA 504
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTATCAAGAG---AAGGAGGAAGAGGAATGGAAAGAAATGAGGAGACCT 561
Qy 181 LysProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
Db 562 AAGCCAAAATTTATCCAGACGAGGAGCGGAGTACACGCCGATCCACCTCAGC 615

RESULT 3
US-09-724-676-12748
; Sequence 12748, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12748

Alignment Scores:
Pred. No.: 3,66e-88 Length: 1931
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12748 (1-1931)
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 93 AGCTCCCTCGATTTGCTGTGGCAACACAGTATCTTACGCGAAGTGAACACCGGCC 152
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 153 AAATTTGAGTCCCTCTTTAGGACGATATGACAGGACATCACCTTTTCAGTATTTTAAAGAC 212
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:09:21 : Search time 52.7555 Seconds  
(without alignments)  
2732.361 Million cell updates/sec

Title: US-09-782-953-3

Sequence: 1 MEVDLDDPSATLACHLDP.....RPKRIQTRRPYPYTHHS 198

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Dgapop 6.0	Dgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.spool/US09782953/runat\_1122002\_114431\_17431/app\_query.fasta\_1.1173  
-DB=Pending Patents NA New -QFMT=fastp -SUFFIX=rnpn -MINMATCH=0 -I -LOCPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR.SCOR=500 -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09782953 @CGN 1 1 70 @unat\_1122002\_114431\_17431 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -NMIT -LONGLOG -DEV.TIMEOUT=120  
-MAIN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Pending\_Patents\_NA\_New\*

1:	/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq*
2:	/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq*
3:	/cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq*
4:	/cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq*
5:	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq*
6:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq*
7:	/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	999.5	95.4	1893	US-09-724-676-12741	Sequence 12741, A
2	999.5	95.4	1893	US-09-724-676A-12741	Sequence 12741, A
3	872.5	83.3	1931	US-09-724-676-12748	Sequence 12748, A
4	872.5	83.3	1931	US-09-724-676A-12748	Sequence 12748, A
5	872.5	83.3	2050	US-09-724-676-12756	Sequence 12756, A
6	872.5	83.3	2050	US-09-724-676A-12756	Sequence 12756, A
7	872.5	83.3	2297	US-09-724-676-12747	Sequence 12747, A
8	872.5	83.3	2297	US-09-724-676A-12747	Sequence 12747, A
9	859.5	82.0	2348	US-10-240-965-60	Sequence 60, Appl
10	847.5	80.9	1839	US-09-724-676-12743	Sequence 12743, A
11	847.5	80.9	1839	US-09-724-676A-12743	Sequence 12743, A

12	847.5	80.9	1876	US-09-724-676-12745	Sequence 12745, A
13	847.5	80.9	1876	US-09-724-676A-12745	Sequence 12745, A
14	731	69.8	1056	US-09-724-676-12742	Sequence 12742, A
15	731	69.8	1056	US-09-724-676A-12742	Sequence 12742, A
16	620.5	59.2	3159	US-10-290-438-1	Sequence 1, Appl1
17	607	57.9	911	US-09-724-676-12749	Sequence 12749, A
18	607	57.9	911	US-09-724-676A-12749	Sequence 12749, A
19	604	57.6	1094	US-09-724-676-12750	Sequence 12750, A
20	604	57.6	1094	US-09-724-676A-12750	Sequence 12750, A
21	604	57.6	1213	US-09-724-676-12740	Sequence 12740, A
22	604	57.6	1213	US-09-724-676A-12740	Sequence 12740, A
23	579	55.2	1002	US-09-724-676-12744	Sequence 12744, A
24	579	55.2	1002	US-09-724-676A-12744	Sequence 12744, A
25	579	55.2	1039	US-09-724-676-12746	Sequence 12746, A
26	579	55.2	1039	US-09-724-676A-12746	Sequence 12746, A
27	578.5	55.2	2411	US-09-724-676-15329	Sequence 15329, A
28	578.5	55.2	2411	US-09-724-676A-15329	Sequence 15329, A
29	539	51.4	615	US-10-290-438-8	Sequence 8, Appl1
30	509.5	48.6	1021	US-10-290-438-3	Sequence 3, Appl1
31	439	41.9	442	US-09-513-999C-1772	Sequence 1772, Ap
32	415.5	39.6	1577	US-09-724-676-12755	Sequence 12755, A
33	415.5	39.6	1577	US-09-724-676A-12755	Sequence 12755, A
34	415.5	39.6	1943	US-09-724-676-12752	Sequence 12752, A
35	415.5	39.6	1943	US-09-724-676A-12752	Sequence 12752, A
36	285.5	27.2	446	US-10-203-138A-2101	Sequence 2101, Ap
37	266.5	25.4	486	US-10-203-138A-827	Sequence 827, App
38	198.5	18.9	531	US-10-290-438-6	Sequence 6, Appl1
39	187.5	17.9	123	US-10-203-138A-5980	Sequence 5980, Ap
40	156	14.9	465	US-09-620-607B-1320	Sequence 1320, Ap
41	150	14.3	557	US-09-724-676-12753	Sequence 12753, A
42	150	14.3	557	US-09-724-676A-12753	Sequence 12753, A
43	147	14.0	740	US-09-724-676A-12754	Sequence 12754, A
44	147	14.0	740	US-09-724-676A-12754	Sequence 12754, A
45	140	13.4	85	US-10-203-138A-7216	Sequence 7216, Ap

#### ALIGNMENTS

RESULT 1  
US-09-724-676-12741  
Sequence 12741, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: CompuGen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 CompuGen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12741  
LENGTH: 1893  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-724-676-12741

#### Alignment Scores:

Pred. No.: 2,296-102  
Score: 999.50  
Percent Similarity: 97.98%  
Best Local Similarity: 95.96%  
Query Match: 95.37%  
DB: 5  
Length: 1893  
Matches: 190  
Conservative: 4  
Mismatch: 3  
Indels: 1  
Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12741 (1-1893)

QY	1	MetGIGlvaIaAplleuGlnaAplleuProSerAlaThrlAlaCyHlsleuAppro	20
DB	25	ATGAGAGAGATGACCTCGAGACCTGCCACGACCATCGCTGACCTGACCGCG	84
QY	21	ArgvalPhvaIaAplleuGlnaAplleuProSerAlaThrlAlaCyHlsleuAppro	40
DB	85	CGGAGTGTGAGACCGCTGCGCGCGCAATTGATTCCTTTAGACGATGAC	144

**THIS PAGE BLANK (usr10)**

```
Db 364 CAGAAAAT---GGCTACGAGAAAAAAGATTCCAGATGCACCCCTCTCTTGATGCC 420
Qy 130 AspleuleuTyraIaIleSerLysIleuGIyProGIyGIuIySTyGIuLeuHleAlaIa 149
Db 421 AGCTCACTGGAGATATCTGTGTCACAGGA-----CAGTATGATTTGGTGAATGC 471
Qy 150 ThrAspProThrProSerValValIvalHlValCyseGIuSerAspGIuInGIuInGIu 169
Db 472 CTGGCTTCCT-----ATTCGAGACCCCAAGACCGAGATGCACATGAT 516
Qy 170 GIuGIuGIuGIu-----MetGIu 175
Db 517 ATTGAGAGACAGTGTCTAGGAGATGGCTGTCTGTGCATCTCACACTATGCCATGATGAAG 576
Qy 176 ArgMetLysArgProLys---ProLysIleIleGIuInThrArgProGIuTyThrPro 194
Db 577 AAGATGCAGTTGCCAGAACTGCCCAAGACATCAGCTACAAAGCA-----TATATTCCA 630
```

Search completed: December 14, 2002, 22:30:09  
Job time : 43.0506 secs





```
Db 1507 GATCATTTATCAGCGATCATGCTTGTCTTCTCTCTGTCTCAGAACTATACGA--- 1451
Qy 39 TyrAspLysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPhe 58
Db 1450 ---GAAACCCCTAGCAGCATGCAATACATCTCTCTTGTGAAACCTTATCTGAGAGGCTCT 1394
Qy 59 SerAsnProLeuSerAlaAlaAspAlaArgLeuLeu---HisLysThrGluPheLeu 77
Db 1393 GTTGAGGTATATACATCCAGTGACGGCTTGAGTCGATGGCCCAAAACA----- 1343
Qy 78 GlyLysGluMetLysLeuTyrPhe----- 85
Db 1342 ---AARGAACTCTCCATCTTCTCAAAAGTTCTTAACACACTTGGAGAGACATTGTATCC 1286
Qy 86 ----- 109
Db 1285 AATGAGCCTTAGTGTGGAAGCCATAGCTGTGTCATCCACATCTTCCACATTCGATT 1226
Qy 94 SerHisLeuAlaProAsnProAspLysGlnPheLeuLeuSerPro----- 109
Db 1225 CCT-----AGCCAGCAAAATCCC-----ATCTTCAGTCCAGTGCTGCTTTCAC 1184
Qy 110 -----ProAlaSerProProValGlyTrpLys 118
Db 1183 ATAGTCCATGCACTGCTCAATCTCTCGTTGGAAGTAGCGGAGATCCC----- 1136
Qy 119 GlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeu 138
Db 1135 -----GAGTCGCTCAACCGATCCACACCAGATGTCTCGAAAGATCCACCGGATA 1082
Qy 139 GlyProGlyGluLys-----TyrGluLeuHis 147
Db 1081 AACATTGGGGAGCTCCCGCTTGAATTTTGTGAAATCTCTGTGATGTATTCGAAGCACTT 1022
Qy 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167
Db 1021 CTTGTCACCGGT-----TTGCATACGCGGTAAAGCGGTAGCCGAGG 980
Qy 168 GluGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThr 187
Db 979 AGAGAACAGAGAGAGCGCTCGCAGGACTGGAGGTTCAGAGCTCGGCCAGTCCAGGTC 920
Qy 188 ArgArgPro-----GluTyrThrProIleHis 196
Db 919 AGGCATCCCTTCAAGCGTGTGAGGATGAGCGTGGGACCTGTGCAT 872
```

## RESULT 12

US-08-405-254-7/c  
Sequence 7, Application US/08405254  
Patent No. 5773288

## GENERAL INFORMATION:

APPLICANT: BRIGGS, STEVEN P.  
APPLICANT: BENSON, ROBERT J.  
TITLE OF INVENTION: PLANT GENES AFFECTING GIBBERELLIC ACID  
TITLE OF INVENTION: BIOSYNTHESIS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: FOLEY & LARDNER  
STREET: 3000 K STREET, NW, SUITE 500  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20007  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,254  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/261,465  
FILING DATE: 17-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 33229/299/PIHI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2784 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-405-254-7

Alignment Scores:  
Pred. No.: 7.75 Length: 2784  
Score: 79.00 Matches: 49  
Percent Similarity: 31.78% Conservative: 26  
Best Local Similarity: 20.76% Mismatches: 79  
Query Match: 7.54% Indels: 82  
DB: 1 Gaps: 11

US-09-782-953-3 (1-198) x US-08-405-254-7 (1-2784)

```
Qy 19 AspProArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThr 38
Db 1507 GATCATTTATCAGGATCATGCTTGTCTTCTCTCTGTCTCAGAACTATACGA--- 1451
Qy 39 TyrAspLysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPhe 58
Db 1450 ---GAAACCCCTAGCAGCATGCAATACATCTCTCTTGTGAAACCTTATCTGAGAGGCTCT 1394
Qy 59 SerAsnProLeuSerAlaAlaAspAlaArgLeuLeu---HisLysThrGluPheLeu 77
Db 1393 GTTGAGGTATATACATCCAGTGACGGCTTGAGTCGATGGCCCAAAACA----- 1343
Qy 78 GlyLysGluMetLysLeuTyrPhe----- 85
Db 1342 ---AARGAACTCTCCATCTTCTCAAAAGTTCTTAACACACTTGGAGAGACATTGTATCC 1286
Qy 86 -----AlaGlnThrLeuHisIleGlySer 93
Db 1285 ATGTAGCCTTAGTGTGGAAGCCATAGCTGTGTCATCCACATCTTTCACATTCGATT 1226
Qy 94 SerHisLeuAlaProAsnProAspLysGlnPheLeuLeuSerPro----- 109
Db 1225 CCT-----AGCCAGCAAAATCCC-----ATCTTCAGTCCAGTGTGCTTTCAC 1184
Qy 110 -----ProAlaSerProProValGlyTrpLys 118
Db 1183 ATAGTCCATGCACTGCTCAATCTCTCGTTGGAAGTAGCGGAGATCCC----- 1136
Qy 119 GlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeu 138
Db 1135 -----GAGTCGCTCAACCGATCCACACCAGATGTCTCGAAAGATCCACCGGATA 1082
Qy 139 GlyProGlyGluLys-----TyrGluLeuHis 147
Db 1081 AACATTGGGGAGCTCCCGCTTGAATTTTGTGAAATCTCTGTGATGTATTCGAAGCACTT 1022
Qy 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167
Db 1021 CTTGTCACCGGT-----TTGCATACGCGGTAAAGCGGTAGCCGAGG 980
Qy 168 GluGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThr 187
Db 979 AGAGAACAGAGAGAGCGCTCGCAGGACTGGAGGTTCAGAGCTCGGCCAGTCCAGGTC 920
Qy 188 ArgArgPro-----GluTyrThrProIleHis 196
```

```

TOPLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
  NAME/KEY: repeat_unit
  LOCATION: 1..407
  OTHER INFORMATION: /rpt_type="terminal"
  FEATURE:
  NAME/KEY: repeat_unit
  LOCATION: 7368..7493
  OTHER INFORMATION: /note="5'UTR"
  NAME/KEY: misc_feature
  LOCATION: 7471..7476
  OTHER INFORMATION: /function="polyA_signal"
  NAME/KEY: repeat_unit
  LOCATION: 7368..7493
  OTHER INFORMATION: /rpt_type="terminal"
  OTHER INFORMATION: /note="3'UTR"
  NAME/KEY: misc_feature
  LOCATION: 408..7367
  OTHER INFORMATION: /product="Coagulation Factor VIII"
  PUBLICATION INFORMATION:
  AUTHORS: Elder, F.
  AUTHORS: Lakich, D.
  AUTHORS: Gitschler, J.
  TITLE: Sequence of the Murine Factor VIII cDNA.
  JOURNAL: Genomics
  VOLUME: 16
  PAGES: 374-379
  DATE: 1993
  RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 7476
PCT-US94-13200-5

Alignment Scores:
Pred. No.: 29.4      Length: 7493
Score: 79.50      Matches: 56
Percent Similarity: 39.19%      Conservative: 31
Best Local Similarity: 25.23%      Mismatches: 83
Query Match: 7.59%      Indels: 52
DB: 5      Gaps: 10

US-09-782-953-3 (1-198) x PCT-US94-13200-5 (1-7493)
QY 15 AAlaCyHlIeUaApRoArG-----ValPheValAepGlyLeuCyArGAlaLys 31
DB 188 GCCTGCTTCTACTACCAAAAGAAAGTATCTTCAGATCTGTTTGTCTAATGCTAC 247
QY 32 PheGluSer-----LeuPheArGThrTyraP 40
DB 248 TTTCACCTCACAGATAACTCCAGAAATCTCTGCAGAAATATTAGGACTTTTAC 307
QY 41 LyAaPThrThrPheGlnTyrrPheLys-----SerPheLysArgValaGlnIleAsn 57
DB 308 TAAATCATTCATTTCTTTGTTGTTTAAAGCTAAAGTATTTAGAGAGAGTTAAAT 367
QY 58 -----PheSerAsnProLeuSerAlaIlaAspAlaArgLeuArg 70
DB 368 TTTCATTTCTTATGTAACATTTTATAGTAATAAAGCATGCAATATGCACTCTTCCG 427
QY 71 LeuHlAlyThrGlu-----PheLeu-GlyLyGluMetLysLeuTyrrPheAlaG 87
DB 428 TTGCTTCTTTCGAGCCTTTTCAATTTCTGCTCTAGTCAGATCAGAAAGTACTCTGG 487
QY 87 nhrLys-----HisIleGlySerSerHlAlyLeuAlaProPheAsnProAs 102
DB 488 TGCAGTGAATGCTCTGGAACATATATTCAGATGATCTGCTCAGTCTGCTGATCAGA 547
QY 102 pLyGlnPheLeuIleSerProProAlaSerProProValaGlyTrpLysGlnValaGlnLys 122
DB 548 CTCAGATTTCTTCTTGAATGTCAGATCTTTTCCATTTCAACACCTCCATCATGTATTA 607

```

```

QY 122 palatnProValIleAsnTyrr---AspLeuTyrrAlaIleSerLys----- 137
DB 608 AAAGACTGTGTGTGAGATACAGACCAAGCCTTTTCAACATTTGCCAACGCCACCC 667
QY 138 -----LeuGlyProGlyGluLysTyrrGluLeuHlS----- 147
DB 668 CTGATGGGTGCTTCAAGTCTTACCATTTTGACTGAGGTTTCATGACACAGTGTACTTAC 727
QY 148 -----AlaAlaThrAspProThrProSerValValaHlSValCyGlu---Se 163
DB 728 ACTTAAACATGCGCTTTCATCTCTGCTGCTTCACTGCTGTGCTGCTACTGGA 787
QY 163 rAspGlnGluAaGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 183
DB 788 AGCTTCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
QY 183 glie 184
DB 848 AGTT 851

RESULT 11
US-08-261-465-3/C
Sequence 3, Application US/08261465
Patent No. 5612191
GENERAL INFORMATION:
APPLICANT: BRIGGS, Steven P.
APPLICANT: BENSEN, Robert J.
TITLE OF INVENTION: PLANT GENES AFFECTING GIBBERELLIC ACID
TITLE OF INVENTION: BIOSYNTHESIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,465
FILING DATE: 17-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/213/PTRI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2784 base pairs
TYPE: nucleic acid
STRANDNESS: double
TOPOLOGY: linear
US-08-261-465-3

Alignment Scores:
Pred. No.: 7.75      Length: 2784
Score: 79.00      Matches: 49
Percent Similarity: 31.78%      Conservative: 26
Best Local Similarity: 20.76%      Mismatches: 79
Query Match: 7.54%      Indels: 82
DB: 1      Gaps: 11

US-09-782-953-3 (1-198) x US-08-261-465-3 (1-2784)
QY 19 AspProArgValaPheValaAspGlyLeuCyArGAlaLysPheGluSerLeuPheArGThr 38

```





```

DB 428 TTGCTTCTTCTGAGCTTTTCAATTTCTGCTAGTCCATGACAGAGATACTACTTGG 487
QY 87 nThrLeu-----HisIleGlySerSerHisLeuAlaProProAsnProAs 102
DB 488 TGCAGTGAATTGCTCTGAACTATATTTCAGAGATGATCTGTGCTGTGCTGATACAGA 547
QY 102 pLySgInPheLeuIleSerProProAlaSerProProValGlyTTrpLySgInValGluAs 122
DB 548 CTGAAGATTTCTTCTTGAATGTCATACATCTTTTCCATTGACACCTTCATCATGTATTA 607
QY 122 pAlaThrProValIleLeuTyr---AspLeuLeuTyrAlaIleSerHis----- 137
DB 608 AAGAGCTGTGTGTGTAGAGTACAGAGCCAGCTTTTCAACATTTGCCAAGCCAGCCACC 667
QY 138 -----LeuGlyProGlyGlyTyrGlyLeuHis----- 147
DB 668 CTGAGTGGTGTGTGTCTAGCTCTACCATTTGGACTGAGCTTACAGACAGTGGTCACTTAC 727
QY 148 -----AlaAlaThrAspProThrProSerValValAlaHisValCysGlu---Se 163
DB 728 ACTTAAACATGAGCTTCTCATCTCTGACGCTTTCATGCTGTGTGTGTCTTACTGAA 787
QY 163 rAspGInGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 183
DB 788 AGCTTCTGAGGAGATGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 847
QY 183 rIle 184
DB 848 AGTT 851

RESULT 8
US-09-037-601-5
Sequence 5, Application US/09037601
Patent No. 6180371
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13300
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Peifer, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7493 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Mus musculus
FEATURE:
NAME/KEY: repeat_unit
LOCATION: 1..407
OTHER INFORMATION: /rpt_type="terminal"
OTHER INFORMATION: /note="5' UTR"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7471..7476
OTHER INFORMATION: /function="polyA signal"
FEATURE:
NAME/KEY: repeat_unit
LOCATION: 7368..7493
OTHER INFORMATION: /rpt_type="terminal"
OTHER INFORMATION: /note="3' UTR"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 408..7367
OTHER INFORMATION: /product="coagulation factor VIII"
PUBLICATION INFORMATION:
AUTHORS: Laskich, F.
AUTHORS: Gitshier, J.
TITLE: Sequence of the murine Factor VIII cDNA
Patent No. 6180371
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
US-09-037-601-5

Alignment Scores:
Pred. No.: 29.4 Length: 7493
Score: 79.50 Matches: 56
Percent Similarity: 39.18% Conservative: 31
Best Local Similarity: 25.23% Mismatches: 83
Query Match: 7.59% Indels: 52
DB: 4 Gaps: 10

US-09-782-953-3 (1-198) x US-09-037-601-5 (1-7493)
QY 15 AlaCyHisLeuAspProArg-----ValPheValAspGlyLeuCyAsnGlnAlaLys 31
DB 188 GCGTCTTCTCTGAGTACCCAAAGAAAGTAATCCTTCAGATGTTTGTGCTAATGCTAC 247
QY 32 PheGluSer-----LeuPheArgThrTyrAsp 40
DB 248 TTTCACCTCAGAGTAATACTCCAGAAATCCTTCGAAATATTTTGGACTTTTAC 307
QY 41 LyAspThrThrPheGlnTyrPheLys-----SerPheLysArgValArgIleAsn 57
DB 308 TAAATCATTAACATTTCTTTTGTCTTAAACCTAAAGTTATTTTGAAGAGCTTAAT 367
QY 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
DB 368 TTTCATTTCTTAGTTGAACATTTTCTAGTATATAAAGCATGCAATGCACTCTTGGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetCysLeuTyrPheAlaGln 87
DB 428 TTGCTTCTTCTGAGCTTTTCAATTTCTGCTCTAGTGCATCAGAAAGATACCTTGG 487
QY 87 nThrLeu-----HisIleGlySerSerHisLeuAlaProProAsnProAs 102
DB 488 TGCAGTGAATTGCTCTGAACTATATTTCAGAGTATGCTGTGCTGTGCTGATACAGA 547
QY 102 pLySgInPheLeuIleSerProProAlaSerProProValGlyTTrpLySgInValGluAs 122

```

```
QY 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 368 TTTCAATCTTTAGTTGAACATTTTCTAGTAATAAAGCCATGCAATAGCACTCTTGGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetLysLeuTyPheAlaG1 87
Db 428 TTGCTTCTTTCTGAGCCTTTTCAATTTCTCTCTAGTGCCATCAGAAGACTACTACCTTGG 487
QY 87 nThrLeu-----HisLeGlySerSerHisLeuAlaProProAsnProAs 102
Db 488 TGCAGTGGAAATGTCCTGGAACTATTTACAGTGCATCTGCTAGTGTGCTGCATACAGA 547
QY 102 pLysGlnPheLeuLeuSerProProAlaSerProProValGlyTrpLysGlnValGluAs 122
Db 548 CTCAGAGATTTCTTCTAGATCAAGGACCACTTTTCCATTCACACCTCCCATCATGTATAA 607
QY 122 pAlaThrProValLeuAsnTy-----AspLeuLeuTyTyAlaIleSerLys----- 137
Db 608 AAGACTGTGTGTTGTAGAGTACAGGACCACTTTTCAACATTTGCCAAGCCCGAGCCACC 667
QY 138 -----LeuGlyProGlyGluLysTyTyGluLeuHis----- 147
Db 668 CTGATGGGTTTGTAGTCTTCTACATTTGGACTGGCTTCATGACACAGTGGTCATTAC 727
QY 148 -----AlaAlaThrAspProThrProSerValValValHisValCysGlu-----Se 163
Db 728 ACTTAAACATGCTTCTCATCTGTCAGTCTTTCATGCTGTTGTTGTTGTTGTTGTTGTTG 787
QY 163 rAspGlnGluAsnGluGluGluGluGluGluGluMetGluArgMetLysArgProLysProLy 183
Db 788 AGCTTCTGAGGAGATGAATATAGATCAGACAGCAAGCCAAATGGAAGGAAGATGTATAA 847
QY 183 sile 184
Db 848 AGTT 851

RESULT 7
US-08-670-707A-5
; Sequence 5, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894

REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7493 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: repeat unit
LOCATION: 1..407
OTHER INFORMATION: /rpt_type= "terminal"
OTHER INFORMATION: /note= "5' UTR"
FEATURE:
NAME/KEY: misc feature
LOCATION: 7471..7476
OTHER INFORMATION: /function= "polyA signal"
FEATURE:
NAME/KEY: repeat unit
LOCATION: 7368..7493
OTHER INFORMATION: /rpt_type= "terminal"
OTHER INFORMATION: /note= "3' UTR"
FEATURE:
NAME/KEY: misc feature
LOCATION: 408..7367
OTHER INFORMATION: /product= "coagulation factor VIII"
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Gitschier, J.
TITLE: Sequence of the murine Factor VIII cDNA
Patent No. 5859204
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
US-08-670-707A-5
Alignment Scores:
Pred. No.: 29.4 Length: 7493
Score: 79.50 Matches: 56
Percent Similarity: 39.19% Conservative: 31
Best Local Similarity: 25.23% Mismatches: 83
Query Match: 7.59% Indels: 52
DB: 2 Gaps: 10

US-09-782-953-3 (1-198) x US-08-670-707A-5 (1-7493)
QY 15 AlaCysHisLeuAspProArg-----ValPheValAspGlyLeuCysArgAlaLys 31
Db 188 GCCTGCTTCTCTACTACTACCCAAAGGAAGTAATCTTCAGATCTGTTGTGCTAATGCTAC 247
QY 32 PheGluSer-----LeuPheArgThrTyArg 40
Db 248 TTTCACTCACAGTAGATAAACTTCCAGAAATCTCTGCAAAATATTTAGGACTTTTAC 307
QY 41 LysAspThrThrPheGlnTyPheLys-----SerPheLysArgValArgIleAsn 57
Db 308 TAAATCATTAACATTTCTTTTGTCTTAAAGCTAAAGTTATTTTGGAGAGAGAGTTAAAT 367
QY 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 368 TTTCAATCTTTTAGTTGAAACATTTTCTAGTAATAAAGCCATGCAATAGCACTCTTGGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetLysLeuTyPheAlaG1 87
```



DATE: 1993  
RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 7476

Percent Similarity: 37.02%  
 Best Local Similarity: 23.56%  
 Query Match: 8.73%  
 DB: 5  
 Conservative: 28  
 Mismatches: 74  
 Indels: 57  
 Gaps: 9

US-09-782-953-3 (1-198) x PCT-US94-09752-1 (1-1820)

QY 2 GUGUVALAARLEUGLNAARLEUPROSERALATHTLALCYAHLEUAP----- 19  
 DB 92 GAAGAGCTTCAAGCTTGATATTTCTTGCTGAACCTTACAGACCTTACAGCAGC 151  
 QY 20 -----ProxValPheValer 25  
 DB 152 AGCAATGAGCGCCAGACATAGCTCCATCCAGACGAGATTAAAGATTAATCAGAC 211  
 QY 26 ---GlyLeuCyArgAlaLeuPheGluSerLeuPheArgThyArgPheSerThr 44  
 DB 212 ATGACATGTGCTACGGATGATGGAGAGCTGTTCCGAGTGGCTCCCGAGACCTC 271  
 QY 45 PheGlnTyrPheLeuSerPheLeuValAlaArgIleAsnPheSerAsnProLeuSerAla 64  
 DB 272 TTCGCC-----AGCCAGATGATCGGCTAT 295  
 QY 65 AAlaerAlaArgLeuArgLeuHisIleArgSerHisIleValAlaThrAspProLeuSer 84  
 DB 296 GCCGAT-----CTCATGACAGCATTTTCATCAT-----TTGCTGAT 334  
 QY 85 PheAlaGlnThrLeuHisIleGlySerSerHisIleValAlaProProAsnPro----- 101  
 DB 335 TACCATTCAGTACCTCTTCAGAGCTGCCACGTTTGATGAGCTCCATGAGTCAACGGTG 394  
 QY 102 -----AspLeuGlnPheLeuIleSerProProAlaSer-----Pro 113  
 DB 395 GAACACACATGATGATATCATGTAATGATGCCCTCCGACACAGGACCGCAGC 454  
 QY 114 ProValGlyTyrPheGlnValAlaGlnPheAlaThrProValIleAsnTyrAspLeuThr 133  
 DB 455 TCCGTGCAATTCAGAAAT-----ACCGACTCAAGCGGATCAGTACACGG 502  
 QY 134 AAlaIleSerLeuLeuGlyProGlyGlyTyrGlyLeuHisIleAlaIleThrAspProThr 153  
 DB 503 TCGATTACTGATTAACCCCGCCCACTCTCCACATG-----GCC 544  
 QY 154 ProSerValValIleHisValCyGluSerSerGlnGlnGlnGlnGlnGlnGln 173  
 DB 545 CCCCAAGAAATTCAGCATTCGATGATGAAGATGATGATGAAGAGAGAGAGAGAG 604  
 QY 174 MetGluArgMetLeuArgProLeu 181  
 DB 605 GAAGATGTGCTGAGAAACCAAA 628

# RESULT 3

US-09-234-332-5/c  
 ; Sequence 5, Application US/09234332A  
 ; Patent No. 6087168  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cedars-Sinai Medical Center  
 ; APPLICANT: Michel F. Levesque, M.D.  
 ; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO  
 ; FILE REFERENCE: P07 41494  
 ; CURRENT APPLICATION NUMBER: US/09/234,332A  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: Paeseq for Windows Version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 3138  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: gene  
 ; LOCATION: (0)...(0)

OTHER INFORMATION: ZIC 1 Protein gene; Genbank Accession D76435  
 US-09-234-332-5

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Mismatches:	Indels:	Gaps:
5.2	81.00	3138	37	12	22	5
Percent Similarity:	40.50%					
Best Local Similarity:	30.58%					
Query Match:	7.73%					
DB:	3					

US-09-782-953-3 (1-198) x US-09-234-332-5 (1-3138)

QY 86 AAlaGlnThrLeuHisIleGlySerSerHisIleValAlaThrAspProThrProSerVal-ValValHis 105  
 DB 950 GCCAGCTCGGCGAAGACGGGGTTGAGC---TTGAAGCGCCCATCGCTCCGCGAAGCGG 894  
 QY 106 LeuIleSerProPro-----AlaSerProProValGlyTyrPheGlnVal 120  
 DB 893 TTGATGCCCGACGCCACAGCTCGTTCGCGCAGCTCGCCCGGAGTGTGCGGAGCGG 834  
 QY 121 ---GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLeuGly 139  
 DB 833 CCAAGGTGCTACGCGCGATCGCTGGTAC-----TGGCGG 798  
 QY 140 ProGlyGlyTyrGlyLeuHisIleAlaIleThrAspProThrProSerVal-ValValHis 159  
 DB 797 CCGGCGTCCAGAGACATCGTGTGCTCGGGGCGAGACCGCGCTCCCGCGCGCGCGCA 738  
 QY 159 ValCyGluSerAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179  
 DB 737 CCTCGCACCGCCAGAGGGAATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678  
 QY 179 rgrProLeuProLeuValIleGlnThrArgProGlnTyrThrProIleHisLeuSer 198  
 DB 677 GAGAGAG-----AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 634

# RESULT 4

US-08-111-939-1  
 ; Sequence 1, Application US/08111939  
 ; Patent No. 5460951  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kawai, Shinji  
 ; APPLICANT: Takeshita, Sunao  
 ; APPLICANT: Okazaki, Makoto  
 ; APPLICANT: Amano, Egon  
 ; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 ; STREET: 1300 I Street, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/111,939  
 ; FILING DATE: 26-AUG-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 324033/92  
 ; FILING DATE: 03-DEC-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 230029/92  
 ; FILING DATE: 28-AUG-1992  
 ; ATTORNEY/AGENT INFORMATION:

Sun Dec 15 08:38:14 2002

REGISTRATION NUMBER: 33,778  
REFERENCE/DOCKET NUMBER: U010815-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800  
TELEFAX: (212) 246-8959  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: cDNA for mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: human  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: foetal  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE: gene library of cDNA  
LIBRARY: gene library of cDNA from foetal  
CLONE: BC-17.8-1 and BC-17.8-2  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9  
MAP POSITION: 21q22.1-q22.2  
FEATURE:  
NAME/KEY: cDNA for mRNA, BC-17.8  
LOCATION: 1..2174  
FEATURE:  
NAME/KEY: untranslated 5'  
LOCATION: 1..48  
FEATURE:  
NAME/KEY: coding sequence  
LOCATION: 49..560  
OTHER INFORMATION: Down Syndrome critical  
region 1 (DSCR1)  
FEATURE:  
NAME/KEY: DSCR1  
LOCATION: 1..171  
IDENTIFICATION METHOD: translation of the  
OTHER INFORMATION: - deduced protein  
OTHER INFORMATION: - proline-rich protein domains  
OTHER INFORMATION: - glutamic acid-rich protein domains  
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains  
FEATURE:  
NAME/KEY: untranslated 3'  
LOCATION: 564..2174  
FEATURE:  
NAME/KEY: 2 poly (A)  
LOCATION: 1541..1546 AND 2132..2137  
US-08-665-040-1

Alignment Scores:  
Pred. No.: 6,31e-96 Length: 2174  
Score: 844.50 Matches: 161  
Percent Similarity: 97.63% Conservative: 4  
Best Local Similarity: 95.27% Mismatches: 3  
Query Match: 80.58% Indels: 1  
Gaps: 2  
DB:

US-09-782-953-3 (1-198) x US-08-665-040-1 (1-2174)

QY 30 AlaLysPheGluSerLeuPheArgThrTyAspLysAspThrPheGlnTyrPheLys 49  
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCCACTTTCAGTATTTTAA 117  
QY 50 SerPheLysArgValArgLeuPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69  
Db 118 AGCTTCAACAGGTGAGATTAATTAATTCAGACACCCCTTCTCCGACGACATGCCAGGCTC 177  
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89  
Db 178 CAGCTGCATGAAGTCTGAGTTTCTGGGAAAGAAATGAAGTTATATTTCCTCAGACCTTA 237

QY 90 HisLleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 105  
Db 238 CACATAGGAAGCTCACACCTGGCTCCGCAATCCAGACAGCAGTTTCTGATCTCCCT 297  
QY 110 ProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValLleAsnTyr 129  
Db 298 CCGGCTCTCCGCCAGTGGGATGGAAACAACACTGGAAAGATGGACCCAGTCATAAATAT 357  
QY 130 AspLeuLeuTyrAlaLleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149  
Db 358 GATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTGCACGACG 417  
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169  
Db 418 ACTGACACCACTCCACAGCTGGTGGTCCATGTATGTGAGAGTGTCAAGAG--AAGGAG 474  
QY 170 GluGluGluGluMetGluArgMetLysArgProLysPheLysLleGlnThrArgArg 189  
Db 475 GAAGAAGAGAAATGGAAAGAAATGAGACACCTAGCCAAAATTTATCCAGACGAGGAG 534  
QY 190 ProGluTyrThrProLleHisLeuSer 198  
Db 535 CCGGAGTACACGCGGATCCACCTCAGC 561

RESULT 2

PCT-US94-09752-1  
; Sequence 1, Application PC/TUS9409752  
; GENERAL INFORMATION:  
; APPLICANT: David S. Strayer and Avinash Chander  
; TITLE OF INVENTION: Compositions and Methods for  
; TITLE OF INVENTION: Targeting Cells and Modulating Pulmonary Surfactant Secretion  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: 210 Lake Drive East, Suite 201  
; CITY: Cherry Hill  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08002  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb  
; MEDIUM TYPE: STORAGE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09752  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/176,218  
; FILING DATE: December 30, 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/114,951  
; FILING DATE: August 31, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: JEFF-0042  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (609) 779-2400  
; TELEFAX: (609) 779-8488  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1820  
; TYPE: nucleic acid  
; STRANDEDNESS: single stranded  
; TOPOLOGY: linear  
; PCT-US94-09752-1

Alignment Scores: 0.112 Length: 1820  
Pred. No.: 91.50 Matches: 49  
Score:

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 30.0506 Seconds  
(without alignments)  
2020.659 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEBVDLQDLPSATYACHLDP.....RPRKXIQTRRPETPIHLIS 198

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing filter 45 summaries

Command line parameters:

-MODEL=frame-p2n.model -DEV=xlh  
-Q/cg2\_1/USPTO.spool/US09782953/runat.11122002.114430.17377/app.query.fasta.1.1173  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=tni -MINMATCH=0.1 -DOORGL=0  
-LISTEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USPR=US09782953 @CGN\_1.1.29 @runat.11122002.114430.17377 -NCPU=6 -ICPU=3  
-NO\_XLXRY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMSOFT-120  
-MAEN\_TIMSOFT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cg2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cg2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cg2\_6/prodata/1/ina/5A.COMB.seq:\*  
4: /cg2\_6/prodata/1/ina/5B.COMB.seq:\*  
5: /cg2\_6/prodata/1/ina/PTCS.COMB.seq:\*  
6: /cg2\_6/prodata/1/ina/Backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	844.5	80.6	2174	2	US-08-665-040-1
2	91.5	8.7	1820	5	PCT-US94-09752.1
3	81	7.7	3138	3	US-09-234-332-5
4	81	7.7	3728	1	US-08-111-939-1
5	79.5	7.6	7493	1	US-08-212-133A-7
6	79.5	7.6	7493	1	US-08-474-503-5
7	79.5	7.6	7493	2	US-08-670-707A-5
8	79.5	7.6	7493	4	US-08-037-601-5
9	79.5	7.6	7493	4	US-08-315-179-5
10	79.5	7.6	7493	5	PCT-US94-13400-5
11	79	7.5	2784	5	US-08-261-465-3
12	79	7.5	2784	1	US-08-405-254-7

13	79	7.5	3429	1	US-08-097-997A-10	Sequence 10, Appl
14	79	7.5	3429	3	US-08-665-574C-10	Sequence 10, Appl
15	79	7.5	3429	4	US-08-946-994-10	Sequence 10, Appl
16	79	7.5	4234	1	US-08-446-038B-1	Sequence 1, Appl
17	79	7.5	4234	1	US-08-446-010B-1	Sequence 1, Appl
18	79	7.5	4234	1	US-08-805-445-1	Sequence 1, Appl
19	79	7.5	4234	2	US-08-064-067D-1	Sequence 1, Appl
20	79	7.5	4234	2	US-09-066-208-1	Sequence 1, Appl
21	79	7.5	36651	4	US-09-738-894A-3	Sequence 3, Appl
22	78.5	7.5	1644	4	US-09-877-730-25	Sequence 25, Appl
23	78.5	7.5	1877	4	US-09-877-730-23	Sequence 23, Appl
24	78.5	7.5	1877	4	US-09-877-730-22	Sequence 22, Appl
25	78.5	7.5	2139	4	US-09-877-730-21	Sequence 21, Appl
26	78.5	7.5	2139	4	US-09-877-730-20	Sequence 20, Appl
27	78.5	7.5	2481	4	US-09-877-730-19	Sequence 19, Appl
28	78.5	7.5	2724	4	US-09-877-730-15	Sequence 15, Appl
29	78.5	7.5	2724	4	US-09-877-730-14	Sequence 14, Appl
30	78.5	7.5	2958	4	US-09-877-730-9	Sequence 9, Appl
31	78.5	7.5	2976	4	US-09-877-730-11	Sequence 11, Appl
32	78.5	7.5	3210	4	US-09-877-730-1	Sequence 1, Appl
33	78.5	7.5	3219	4	US-09-877-730-17	Sequence 17, Appl
34	78.5	7.5	3453	4	US-09-877-730-7	Sequence 7, Appl
35	78.5	7.5	3874	4	US-09-877-730-31	Sequence 31, Appl
36	78.5	7.5	6171	1	US-08-459-568-1	Sequence 1, Appl
37	78.5	7.5	6171	2	US-08-399-411-1	Sequence 1, Appl
38	78.5	7.5	6171	3	US-08-516-859A-1	Sequence 1, Appl
39	78.5	7.5	6171	4	US-09-586-472-1	Sequence 1, Appl
40	78.5	7.5	6171	4	US-09-528-706-1	Sequence 1, Appl
41	78	7.4	1913	2	US-08-588-258B-41	Sequence 41, Appl
42	78	7.4	1913	5	PCT-US96-08295-41	Sequence 6, Appl
43	77	7.3	1086	2	US-08-985-090-6	Sequence 3, Appl
44	77	7.3	1086	3	US-09-165-543-33	Sequence 3, Appl
45	77	7.3	1740	4	US-08-796-101-50	Sequence 50, Appl

#### ALIGNMENTS

RESULT 1  
US-08-665-040-1  
Sequence 1, Application US/08665040  
Patent No. 5869318  
GENERAL INFORMATION:  
APPLICANT: ESTIVILL, PALLEJA, XAVIER  
INVENTOR: PRITCHARD, MELANIE  
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,  
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING PCR" TECHNIQUE,  
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCTRI) HIGHLY  
TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD  
FOR CHARACTERIZING IT.  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PARRY  
STREET: 26 WEST 61ST STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1 FOR DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,040  
FILING DATE: JUNE 7, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: ES 89501140  
FILING DATE: JUNE 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: JANET I. CORD



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 14, 2002, 20:02:01 ; Search time 1822.74 Seconds  
(without alignments)  
2731.177 Million cell updates/sec

Title: US-09-782-953-3  
Perfect score: 1048  
Sequence: 1 MEEVDLQDLPATSIACHLDP.....RPKXIIQTRRDEYTPIHLS 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues  
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.epool/US09782953/runat\_11122002\_114430\_17408/app\_query.fasta.1.1173  
-DB=Pending Patents NA Main -QMT=fastap -SUFFIX=nrnp -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bite -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09782953 -CGN 1 1 2457 -orunat\_11122002\_114430\_17408 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAPP -LARGESQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:  
1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq.\*

29: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US099D\_COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US100A\_COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US100B\_COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US101A\_COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US101B\_COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US102A\_COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US102B\_COMB.seq.\*  
44: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq.\*  
45: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq.\*  
46: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq.\*  
47: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq.\*  
48: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq.\*  
49: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq.\*  
50: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq.\*  
51: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq.\*  
52: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq.\*  
53: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq.\*  
54: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq.\*  
55: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq.\*  
56: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq.\*  
57: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq.\*  
58: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq.\*  
59: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq.\*  
60: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq.\*  
61: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq.\*  
62: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq.\*  
63: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq.\*  
64: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq.\*  
65: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq.\*  
66: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq.\*  
67: /cgn2\_6/ptodata/1/pna/US6023\_COMB.seq.\*  
68: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq.\*  
69: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq.\*  
70: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq.\*  
71: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq.\*  
72: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq.\*  
73: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq.\*  
74: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq.\*  
75: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq.\*  
76: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq.\*  
77: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq.\*  
78: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq.\*  
79: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq.\*  
80: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq.\*  
81: /cgn2\_6/ptodata/1/pna/US6037\_COMB.seq.\*  
82: /cgn2\_6/ptodata/1/pna/US6038\_COMB.seq.\*  
83: /cgn2\_6/ptodata/1/pna/US6039\_COMB.seq.\*  
84: /cgn2\_6/ptodata/1/pna/US6040\_COMB.seq.\*  
85: /cgn2\_6/ptodata/1/pna/US6041\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1048	100.0	597	22	US-09-575-580B-2
2	1048	100.0	597	30	US-09-782-953-2
3	1040	99.2	597	80	US-60-360-207-8520
4	1002.5	95.7	2212	30	US-09-782-953-14
5	1002.5	95.7	2407	40	US-10-104-047-922
6	957.5	91.4	2358	1	PCT-US00-05918-347





Db 85 GCGCGTTCGTGGAGGGCGCTGGCCGGGCGAAATTGAGTCCCTCTTAGGACGTATGAC 144

Qy 41 LysAspThrTrpPheGlnTyrPheIysSerPheIysAsnValaGllLeaAsnPheSerAsn 60

Db 145 AAGGACATCACCTTCAGTATATTTTAAAGGCTTCAACACAGTCAGAAATTAACCTTCAGCAAC 204

Qy 61 ProLysSerAlaAlaAspAlaArgLeuArgLeuHisIleValThrGlnPheLeuGlyLeu 80

Db 205 CCGCTTCGCGACGACGAGTCCAGGCTCCAGCTGCATTAAGATGAGTTCCTGGGAAAGAA 264

Qy 81 MetIleuTyrPheAlaGlnThrIleuHisIleGlySerSerHisIleuAlaProProAsn 100

Db 265 ATGAAGTATATATTTGGCTCACACCTTACATATAGAAAGCTCACACCTGGCTCCGCCAAT 324

Qy 101 ProAspIleGlnPheLeuIleSerProProAlaSerProProValGlyTyrIleGlnVal 120

Db 325 CCAGCAAGCAGAGTTTGTGATCTCCCTCCCTCCGCGCAGGGGATGGAACAAAGTG 384

Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerIleLeuGlyPro 140

Db 385 GAAAGTGGACCCCACTATTAACCTATGATCTCTTAATGCGCATCTCCAACTGGGGCCA 444

Qy 141 GlyIleuTyrGlnLeuHisAlaAlaThrAspThrProSerValValHisVal 160

Db 445 GGGGAAAGTATGAAATTCAGCAGCAGCGACTGACACCACTCCGACGCGTGGTGCATGTA 504

Qy 161 CysGlnSerAspGlnGluAsnGlnGluGluGluGluMetGluAlaGlyMetLysArgPro 180

Db 505 TGTAGAGTGTGATCAAGAG--AAAGAGGAAGAAAGAGAAATGGAAGAAATGAGAGAGACT 561

Qy 181 LysPheLysIleIleGlnThrArgArgProGluTyrThrProIleHisIleuSer 198

Db 562 AAGCCAAAATTTATCCAGACCGAGAGGCCGAGATACGCCGATCCACTTAAAG 615

```

RESULT 5
US-10-104-047-922
US-10-104-047-922
; Sequence 922, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10-104, 047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 922
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-922

Alignment Scores:
Pred. No.: 5,25e-91 Length: 2407
Score: 1002.50 Matches: 191
Percent Similarity: 97.99% Conservative: 3
Best Local Similarity: 96.46% Mismatch: 3
Query Match: 95.66% Indels: 1
DB: Gaps: 1

US-09-782-953-3 (1-196) x US-10-104-047-922 (1-2407)

Oy 1 MetClnGlnValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 218 ATGAGAGAGGTGACCTGCGAGACCTGCGCCAGCCGACCATCCCTGTCACCTGAGCCCG 27
Oy 21 ArgValPheValAspGlyLeuGlyCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 278 CGCGTGTCTTCGAGACCGGCTGTGCGGGGCCAAATTGATCCCTCTTTAGAGCATATGAC 33
Oy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

```

Db	338	AAGGACATCACCTTTCAGATATTTTAAGAGCTTCAACAGCTCAGAAATTAACCTTAGCAAC	397
QY	61	ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisIleuYThrGluPheLeuGlyLysGlu	80
Db	398	CCCTTCTCCGCGAGCAGATGCGCAGGCTCCAGCTGCATGAAGTACGATGTTCTTCGGGAAGAA	457
QY	81	MetLysIleuYThrPheAlaGlnThrLeuHisAlileuGlySerSerHisIleuAlaProProAsn	100
Db	458	ATGAAGTTATATTTTGCTCCAGACCTTACACATAGGAAGCTCACACCTCCGCTCCCAAAAT	517
QY	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPheGlnVal	120
Db	518	CCAGCAAGACGATTTCTGATCTTCCCTCCGCTCCGCTCCGCGAGGGATGGAAACAAATG	577
QY	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140
Db	578	GAAAGTGGACCCCGACAGTCATAACTAATGATCTCTTAATGTCATCTCCAAAGCTGGGCCA	633
QY	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal	160
Db	638	GGGGAAAAGTATGAAATTCAGCAGCAGCACTACACCACTCCACAGCGTGTGTCATGTA	697
QY	161	CysGluSerAspGlnGluAsnGluGluGluGluGluGluMetGluValGlyMetLysArgPro	180
Db	698	TGTGAGATGTATCAAGAG--AAGAGAGAAAGAAAGAAATGGAAGAAATGACAGACACT	755
QY	181	LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisIleuSer	198
Db	755	AAGCCAAAATTTATCCAGACCCAGAGGCCGAGATGACCGCCGATCCCACTTAGC	808

```

RESULT 6
PCT-US00-05918-347
: Sequence 347, Application PC/TUS0005918
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben
: TITLE OF INVENTION: Human Lung Cancer Associated Gene Sequences and Polypeptides
: FILE REFERENCE: P1104PCT
: CURRENT APPLICATION NUMBER: PCT/US00/05918
: CURRENT FILING DATE: 2000-03-08
: EARLIER APPLICATION NUMBER: 60/124,270
: EARLIER FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 347
: LENGTH: 2358
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US00-05918-347

Alignment Scores:
Pred. No.: 1,916-86 Length: 2358
Percent Similarity: 95.50 Matches: 188
Best local Similarity: 96.46 Conservative: 3
Query Match: 94.95 Mismatches: 6
DB: 91.64 Indels: 2
Gaps: 1

US-09-782-953-3 (1-198) x PCT-US00-05918-347 (1-2358)
QY 1 MetGluGluValAspLeuGlnAspLeuProSerValaThrIleAlaCysHisLeuAspPro 20
Db 114 ATGAGAGAGGTGGACCTGGACGAGACTGGCCACGCCACCATCGCCCTGTCACTGGACCCG 173
QY 21 ArgValaPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 174 CGCGTGTCTGGACCGGCCCTGTCCGGGCCAAATTTGAGTCCCTCTTTGGACGTATGAC 233
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValaArgIleAsnHisSerAsn 60
Db 234 AAGGACATCCACTTTTACGTAATTTTAAGAGCTTCAACGAGTCAGATTAATCACTTCAGCAAC 293

```

```
Qy 61: ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLeuThrGluPheLeuGlyLysGlu 80
|||
Db 294 CCCTTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 353
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLeuGlySerSerHisLeuAlaProAsn 100
|||
Db 354 ATGAAGTTATATTTTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCGCA-AA 412
|||
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
|||
Db 413 CCAGACAGCAGGTTTCTGATCTCCCTCCCGCTCTCCGSCAGTGGGATGGAACAAGTG 472
|||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
|||
Db 473 GAAGATGCGACCCAGTCATAAATGATGATCTCTTATATGCCATCTCCAAGCTGGGCA 532
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 533 GGGGAAAAGATGATGAATTCACAGCAGGAGCTGACACCACTCCAGCGTGGTCCATGTA 592
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
|||
Db 593 TGTRAGAGTGCATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 649
|||
Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
Db 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACAGCGGATCCACCTCAGC 703
|||
RESULT 7
US-09-925-302-347
; Sequence 347, Application US/09925302
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 347
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-347
Alignment Scores:
Pred. No.: 1,91e-86 Length: 2358
Score: 957.50 Matches: 188
Percent Similarity: 96.46% Conservative: 3
Best Local Similarity: 94.95% Mismatches: 6
Query Match: 91.36% Indels: 2
DB: 34 Gaps: 1
US-09-782-953-3 (1-198) x US-09-925-302-347 (1-2358)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 114 ATGAGGAGGTGGACCTCGAGGACCTGCCAGCGCCACCATCGCTGTCACTGGACCCG 173
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 174 CGCGTGTTCGTGGACGGCTGTGCGGGGCAAAATTTAGTCCCTCTTTAGACGTATGAC 233
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 234 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCAGATAAATCTCAGCAAC 293
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLeuThrGluPheLeuGlyLysGlu 80
|||
```

```
Db 294 CCCTTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 353
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLeuGlySerSerHisLeuAlaProAsn 100
|||
Db 354 ATGAAGTTATATTTTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCGCA-AA 412
|||
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
|||
Db 413 CCAGACAGCAGGTTTCTGATCTCCCTCCCGCTCTCCGSCAGTGGGATGGAACAAGTG 472
|||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
|||
Db 473 GAAGATGCGACCCAGTCATAAATGATGATCTCTTATATGCCATCTCCAAGCTGGGCA 532
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 533 GGGGAAAAGATGATGAATTCACAGCAGGAGCTGACACCACTCCAGCGTGGTCCATGTA 592
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
|||
Db 593 TGTRAGAGTGCATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 649
|||
Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
Db 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACAGCGGATCCACCTCAGC 703
|||
RESULT 8
US-09-644-868-7429
; Sequence 7429, Application US/09644868
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1168-001
; CURRENT APPLICATION NUMBER: US/09/644,868
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,063
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 10075
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7429
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-644-868-7429
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
US-09-782-953-3 (1-198) x US-09-644-868-7429 (1-1615)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 154 ATGAGGAGGTGGACCTCGAGGACCTGCCAGCGCCACCATCGCTGTCCCTGGACCCG 213
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 214 CGCGTGTTCGTGGACGGCTGTGCGGGGCAAAATTTAGTCCCTCTTTAGACGTATGAC 273
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal-ArgIleAsnPheSerAs 60
Db 274 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGCTTCGAATAAATCTCAGCA 333
|||
```

QY 60 nProLeuSer-AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80  
DB 334 CCCCTTCCCGCAG 393  
QY 80 LMetLysLeuLysPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProA 100  
DB 394 AATGAAAGTTATATTTTCTTCAAGCTTACACATAGAGAACTCAACCTGGCTCCGCCAA 453  
QY 100 nProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120  
DB 454 ATCCAGACAGACAGATTCTGATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 513  
QY 120 ALGluAspAlaThrProValAlaSerLysLeuLysGlyTrpLysGlnV 140  
DB 514 TGAAGATGGAGCCAGCTATTAACCTATGATCTTATATGCACTTCCAACTCCAGCTGGGC 573  
QY 140 roGlyGluLysTyrgLysLeuHisAlaAlaThrAspProThrProSerValValHisV 160  
DB 574 CAGGGGAAAAGTATGAAATTCACGCGACCTGACACACTCCACAGCTGGTGCTCATG 633  
QY 160 aLysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 180  
DB 634 TATGTAGAGATGATCAAGAG--AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690  
QY 180 roLysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198  
DB 691 CTAAAGCCAAAATTTATCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746

## RESULT 9

US-09-644-869-8165

Sequence 8165, Application US/09644869

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

APPLICANT: McCarthy, Sean A.

APPLICANT: Holzman, Douglas A.

APPLICANT: Monahan, John

APPLICANT: Richardson, Jennifer

TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

TITLE OF INVENTION: THEREFOR

FILE REFERENCE: 1600.1182-001

CURRENT APPLICATION NUMBER: US/09/644.869

CURRENT FILING DATE: 2000-08-28

PRIOR APPLICATION NUMBER: 60/151,062

PRIOR FILING DATE: 1999-08-27

NUMBER OF SEQ ID NOS: 9708

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8165

LENGTH: 1615

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(1615)

OTHER INFORMATION: n = A,T,C or G

US-09-644-869-8165

## Alignment Scores:

Pred. No.:	3,04e-86	Length:	1615
Score:	953.50	Matches:	188
Percent Similarity:	95.50%	Conservative:	3
Best Local Similarity:	94.00%	Mismatches:	6
Query Match:	90.98%	Indels:	3
DB:	25	Gaps:	1

US-09-782-953-3 (1-198) x US-09-644-869-8165 (1-1615)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
DB 154 ATGAGAGAGGTGAGACCTGACAGAGCTGCCAGCGCACCATCGCTGTGCTGAGACCG 213  
QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyraAsp 40

DB 214 CGCTGTTCGTGAGACGAGCTGTGCGGGCAAAATTGAGTCCCTTTANGACGTATGAC 273  
QY 41 LysAspThrPheGlnTyrrPheLysSerPheLysValGlyValArgIleAsnPheSerAs 60  
DB 274 AAGGACATCACCTTTACAGTNTTTTAAAGCTTCAAAACAAGTCTCGAAATTAACCTCAGCAA 333  
QY 60 nProLeuSer-AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80  
DB 334 CCCCTTCCCGCAG 393  
QY 80 LMetLysLeuLysPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProA 100  
DB 394 AATGAAAGTTATATTTTCTTCAAGCTTACACATAGAGAACTCAACCTGGCTCCGCCAA 453  
QY 100 nProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120  
DB 454 ATCCAGACAGACAGATTCTGATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 513  
QY 120 ALGluAspAlaThrProValAlaSerLysLeuLysGlyTrpLysGlnV 140  
DB 514 TGAAGATGGAGCCAGCTATTAACCTATGATCTTATATGCACTTCCAACTCCAGCTGGGC 573  
QY 140 roGlyGluLysTyrgLysLeuHisAlaAlaThrAspProThrProSerValValHisV 160  
DB 574 CAGGGGAAAAGTATGAAATTCACGCGACCTGACACACTCCACAGCTGGTGCTCATG 633  
QY 160 aLysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 180  
DB 634 TATGTAGAGATGATCAAGAG--AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690  
QY 180 roLysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198  
DB 691 CTAAAGCCAAAATTTATCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746

## RESULT 10

US-09-644-871-7247

Sequence 7247, Application US/09644871

GENERAL INFORMATION:

APPLICANT: Holzman, Douglas A.

APPLICANT: Pan, Yang

TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

TITLE OF INVENTION: THEREFOR

FILE REFERENCE: 1600.1167-001

CURRENT APPLICATION NUMBER: US/09/644.871

CURRENT FILING DATE: 2000-08-28

PRIOR APPLICATION NUMBER: 60/151,059

PRIOR FILING DATE: 1999-08-27

NUMBER OF SEQ ID NOS: 9739

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7247

LENGTH: 1615

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(1615)

OTHER INFORMATION: n = A,T,C or G

US-09-644-871-7247

## Alignment Scores:

Pred. No.:	3,04e-86	Length:	1615
Score:	953.50	Matches:	188
Percent Similarity:	95.50%	Conservative:	3
Best Local Similarity:	94.00%	Mismatches:	6
Query Match:	90.98%	Indels:	3
DB:	25	Gaps:	1

US-09-782-953-3 (1-198) x US-09-644-871-7247 (1-1615)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20  
DB 154 ATGAGAGAGGTGAGACCTGACAGAGCTGCCAGCGCACCATCGCTGTGCTGAGACCG 213

```
QY 21 ArgValPheValAspGlyLeuCyseArgAlaLysPheGluSerLeuPheArgThrTyAsp 40
|||
DB 214 CGCGTGTTCGTGAGCGCCCTGTCGGGCCAAATTTGAGTCCCTCTTTANGACGTATGAC 273

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal - ArgIleAsnPheSerAs 60
|||
DB 274 AAGGACATCACCTTTGAGTNTTTAAGAGCTTCAACCGAGCTTCGAATAAATCTTCAGCAA 333

QY 60 nProLeuSer - AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
|||
DB 334 CCCCTTCTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGG 393

QY 80 LuMetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProA 100
|||
DB 394 AAATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 453

QY 100 enProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120
|||
DB 454 ATCCAGACAAGCAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAG 513

QY 120 alGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyP 140
|||
DB 514 TGAAGATGCCGCCAGCTATGAATTCACGACGAGCTGACACCACTCCAGCGTGGTCCATG 573

QY 140 roGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisV 160
|||
DB 574 CAGGGGAAAAGTATGAATTCACGACGAGCTGACACCACTCCAGCGTGGTCCATG 633

QY 160 alCy6GluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgP 180
|||
DB 634 TATGTGAGTGTATCAAGAG --- AAGGAGGAAGAAGAGGAAATGGAAAGATGAGGAGAC 690

QY 180 roLysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
DB 691 CTAAGCCAAAATTTATCCAGACCAAGAGGCCGAGTACACGCCGATCCACCTCAGC 746
```

## RESULT 11

```
US-09-652-123-7916
; Sequence 7916, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652.123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7916
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-123-7916
```

```
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
```

US-09-782-953-3 (1-198) x US-09-652-123-7916 (1-1615)

```
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyseHisLeuAspPro 20
|||
DB 154 ATGGAGAGGTGGACCTTGAGGAGCTGCCAGGCCACCATCGCTGTCTGTCGACCCG 213
```

```
QY 21 ArgValPheValAspGlyLeuCyseArgAlaLysPheGluSerLeuPheArgThrTyAsp 40
|||
DB 214 CGCGTGTTCGTGAGCGCCCTGTCGGGCCAAATTTGAGTCCCTCTTTANGACGTATGAC 273

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal - ArgIleAsnPheSerAs 60
|||
DB 274 AAGGACATCACCTTTGAGTNTTTAAGAGCTTCAACCGAGCTTCGAATAAATCTTCAGCAA 333

QY 60 nProLeuSer - AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
|||
DB 334 CCCCTTCTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGG 393

QY 80 LuMetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProA 100
|||
DB 394 AAATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 453

QY 100 enProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120
|||
DB 454 ATCCAGACAAGCAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAG 513

QY 120 alGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyP 140
|||
DB 514 TGAAGATGCCGCCAGCTATGAATTCACGACGAGCTGACACCACTCCAGCGTGGTCCATG 573

QY 140 roGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisV 160
|||
DB 574 CAGGGGAAAAGTATGAATTCACGACGAGCTGACACCACTCCAGCGTGGTCCATG 633

QY 160 alCy6GluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgP 180
|||
DB 634 TATGTGAGTGTATCAAGAG --- AAGGAGGAAGAAGAGGAAATGGAAAGATGAGGAGAC 690

QY 180 roLysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
DB 691 CTAAGCCAAAATTTATCCAGACCAAGAGGCCGAGTACACGCCGATCCACCTCAGC 746
```

## RESULT 12

```
US-09-652-915-9010
; Sequence 9010, Application US/09652915
; GENERAL INFORMATION:
; APPLICANT: Falb, Dean R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1175-001
; CURRENT APPLICATION NUMBER: US/09/652.915
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,110
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 10311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9010
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-915-9010
```

```
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
```

US-09-782-953-3 (1-198) x US-09-652-915-9010 (1-1615)

```
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyseHisLeuAspPro 20
|||
```



```

Db 154 ATGAGAGAGTGGACCTGGACAGACTCCGACGCCACCATGGCTGTCTGACCCG 213
Qy 21 ArgValPheValIlePglYLeuCyArgAlaYaphneGluSerLeuPheArgThrTyArg 40
Db 214 CGCGTGTCTGGACGGCGCTGTGCGGGCCAAATTGAGTCCCTCTTANGAGCTATGAC 273
Qy 41 LysAspThrThrPheGlnTyPheYsSerPheYsArgVal-ArgIleAsnPhseSerA 60
Db 274 AAGGACATCCACCTTCACTGTTTAAAGCTTCAAGACTTCGATAAATTCAGCAA 333
Qy 60 nProLeuSer-AlaIleAspAlaArgLeuArgLeuHisYsThrGluPheLeuGlyLys 80
Db 334 CCGCTTTCCTCCGACGAGATGCCAGGCTCCAGCTCAAGACTGAGTTTGGGAAAG 393
Qy 80 MetLeuLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPro 100
Db 394 AATGAAATTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCC 453
Qy 100 nProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGln 120
Db 454 ATCCAGCAAGCAGTTCTCATCTCCCTCCCTCCGCTCCGACATGGGAAACAAG 513
Qy 120 AlGluAspAlaThrProValIleAsnTyAspLeuLeuTyAlaIleSerLysLeuGly 140
Db 514 TGGAGATGGACACCCGACATATACTATGATCTTTATGCTCATCTCCAGCTGGG 573
Qy 140 GGIyGluYsTyPheGluLeuHisAlaIleThrAspProThrProSerValIleHisVal 160
Db 574 CAGGGGAAAAGTATGATTCAGCAGCAGCAGCTCACACACTCCGACGGTGGTCCAG 633
Qy 160 AlCyGluSerAspGlnGluAsnGluGluGluGluMetGluAspGlyAsp 180
Db 634 TAGTGAAGTGAATCAAGAG--AAGGAGCAAGAAAGAAATGGAAGAAAGAGAGAC 690
Qy 180 LysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
Db 691 CTAAAGCAAAATATATCCAGACAGAGAGCCGAGATACAGCCGATCCACTCAGC 746

```

## RESULT 13

```

US-09-782-953-5
Sequence 5, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-09-782-953-5

```

## Alignment Scores:

```

Pred. No.: 1,99e-83 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 30 Gaps: 1

```

US-09-782-953-3 (1-198) x US-09-782-953-5 (1-597)

```

Qy 4 ValAspLeuGlnAspLeuPro-----SerIleThrIleAlaCyHisIleAspPro 20
Db 1 ATGATTTTATAGGACTTTATAGCTACAAATTTTACCTCCGATTTGCTGTGGCAAAAGCAT 60
Qy 21 ArgValPheValIlePglYLeuCyArgAlaYaphneGluSerLeuPheArgThrTyArg 40
Db 61 GATGCTTCAGCGAAAAGTGAAGCAGGCGCAAAATTGAAATCCCTCTTCAGAAACATATGAC 120
Qy 41 LysAspThrThrPheGlnTyPheYsSerPheYsArgVal-ArgIleAsnPhseSerA 60
Db 121 AAGGACATCCACCTTCACTGTTTAAAGCTTCAAGACTTCGATAAATTCAGCAA 180
Qy 61 ProLeuSerAlaIleAspAlaArgLeuArgLeuHisYsThrGluPheLeuGlyLys 80
Db 181 CCGTATATCGACGCGAGTCCAGGCTGGCTGCACAAAGCCGATTCCTGGGAAAGGAA 240
Qy 81 MetLeuLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 241 ATGAAATTGTATTTTGTCTCAGACCTTACACATAGGAAGTTACACACTGGCTCCGCCAAT 300
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGlnVal 120
Db 301 CCGACAAACGTTCTCATCTCCCTCCGCTCTCTCCGCTCCGACATGGGAAACAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyAspLeuLeuTyAlaIleSerLysLeuGlyPro 140
Db 361 GAAATGCCACCCCGCATATAATTACATCTTTATGCTCATCTCCAGCTGGG 420
Qy 141 GGIyGluYsTyPheGluLeuHisAlaIleThrAspProThrProSerValIleHisVal 160
Db 421 GGAAGAAAGTATGAACTGATGTCAGCAGCAGACCCCACTCCAGTGGTGGTCCACCTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluAspGlyAspPro 180
Db 481 TGTGAAGTGAATCAAGAG--AAGGAGCAAGAAAGAAATGGAAGAAAGAGAGACCC 540
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
Db 541 AAGCCCAAAATATATCCAGACAGAGAGCCGAGATACAGCCGATCCACTTAC 594

```

## RESULT 14

```

US-09-782-953-11
Sequence 11, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 2331
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (144)..(734)
US-09-782-953-11

```

## Alignment Scores:

```

Pred. No.: 8,16e-78 Length: 2331
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 30 Gaps: 1

```

US-09-782-953-3 (1-198) x US-09-782-953-11 (1-2331)





**THIS PAGE BLANK (use 70)**